Neural Networks Applied to Analytical Procedures

A thesis submitted in partial completion of the degree of Master of Commerce in Accounting

Stewart Li

Supervisors: Richard Fisher and Michael Falta

An Academic Year 2017/2018

Business and Law School, University of Canterbury

ABSTRACT

Both U.S. and international standards require auditors to perform analytical procedures during the audit planning phase to assess the risk of material misstatements in the financial statements and, near the end of the audit, to determine whether the financial statements are consistent with the auditor's understanding of the entity. It is also permissible to use analytical procedures as a substantive procedure. A key step underlying the application of analytical procedures is the formation of a precise expectation, which subsequently affects the effectiveness of analytical procedures. The type of analytical procedure techniques used to develop expectation models is left to the auditor's discretion. According to auditing standards, it can be any technique ranging from simple comparisons of items to sophisticated analytical models. However, in a data-rich environment, the effectiveness of traditional, simple, analytical procedure methods has been recently questioned, and more advanced approaches have been called for. Among modern statistical and machine-learning methods, neural networks have proved to be useful in both pattern recognition and prediction. Coakley and Brown's (1993) study was the first to research the application of neural networks as an analytical procedure to direct auditors' attention. Following their recommendations, the current study extends their work by incorporating input data obtained from both audited, periodic financial statements of multiple firms and exogenous variables to study analytical procedure techniques of varying levels of sophistication. This study used an experimental design to examine the relative effectiveness of two well-documented analytical review techniques (ratio analysis and regression analysis) and an alternative approach, artificial neural networks. Archival data were obtained from seven listed Chinese companies operating in the dairy industry in order to train and test alternative techniques. The methodology for the study is discussed in detail in five subsections. Results suggest that the neural network approach was not significantly more effective than financial ratios and regressions, and none of the three approaches provided more overall effectiveness than a purely random procedure. However, the neural network approach did yield considerably fewer Type II errors than the other methods.

Keywords: Analytical procedures, Effectiveness, Neural networks, Regression analysis, Ratio analysis

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to the following individuals who continuously supported me to complete the thesis.

- ✓ My supervisors, Associate Professor Richard Fisher and Associate Professor Michael Falta, for their patient guidance and valuable advice. Otherwise, the thesis would not have been possible.
- ✓ My lovely wife, Song Peng, for her love and encouragement.

DECLARATION

I declare that the thesis which I hereby submit for a Master of Commerce degree at the University of Canterbury is my own work. Relevant materials used here have been duly acknowledged and referenced in accordance with general university requirements.

TABLE OF CONTENTS

ABSTRACT		2
ACKNOWLEDGMENTS		4
DECLARATION		5
LIST OF FIGURES		7
LIST OF TABLES		7
1	**INTRODUCTION**	8
2	**LITERATURE REVIEW**	13
2.1	Analytical procedures	13
2.1.1	Use of analytical procedures	13
2.1.2	Error detection ability of analytical procedures	16
2.1.3	Predictive ability of analytical procedures	18
2.1.4	Analytical procedure techniques	21
2.2	Neural networks	24
3	**HYPOTHESES DEVELOPMENT**	27
4	**METHODOLOGY**	29
4.1	Description of case firms	30
4.2	Definition of sources of error	32
4.3	Investigation rule	35
4.4	Analytical procedure methods	37
4.4.1	Financial ratio procedure	37
4.4.2	Regression procedure	38
4.4.3	Neural network procedure	39
4.4.3.1	The neural network architecture	40
4.4.3.2	Selection of the learning algorithm	43
4.4.3.3	Training process	46
4.4.3.4	The statistical decision rules	48
4.5	Experimental procedures	48
5	**RESULTS**	51
5.1	Financial ratios	51
5.2	Comparison of methods	53
5.3	Additional analyses	56
5.3.1	Effect of error size	57
5.3.2	Effect of statistical level of confidence	59
5.3.3	Effect of source of error	60
5.3.4	Practical application of methods to holdout case	63
6	**DISCUSSION AND LIMITATIONS**	66
7	**CONCLUSIONS**	69
REFERENCES		72
APPENDIX I - LIST OF CASE FIRMS		85
APPENDIX II - COMPARIONS BETWEEN FIRM'S AND AVERAGED RATIOS		86
APPENDIX III - THE USE OF 15, 20 AND 30 TIMES M		89
APPENDIX IV - PARAMETERS OF THE NEURAL NETWORK MODEL		90

LIST OF FIGURES

Figure 1	The process of risk assessment in relation to analytical procedures	15
Figure 2	Averages of financial account balances	34
Figure 3	Financial ratios calculated from audited balances	55
Figure 4	Comparison of analytical procedure methods	56
Figure 5	Effect of alpha risk level on Type I and Type II error rates	58
Figure 6	Effect of alpha risk level on Type I and Type II error rates	60
Figure 7	Effect of source of error on Type I and Type II error rates	62
Figure 8	Regression and neural network methods error rates by individual account	63

LIST OF TABLES

Table 1	Financial accounts included in the study	33
Table 2	Impact of source of error on financial accounts	35
Table 3	Types of attention-directing decisions	37
Table 4	Impact of source of error on evaluated financial ratios	39
Table 5	Parameters and coefficient of regression models	42
Table 6	AIC of neural network analysis	48
Table 7	Comparison of financial ratio analysis	51
Table 8	Comparison of analytical procedure methods	56
Table 9	Effect of error size on error rates	58
Table 10	Effect of statistical level of confidence on error rates	59
Table 11	Effect of source of error on error rates	61
Table 12	Error rates by individual financial account	62
Table 13	The developed models applied for testing	65

1 INTRODUCTION

Financial statements are prepared by audit clients, audited by auditors, and are meant to provide information to investors for decision-making. The primary objective of the external auditor conducting an audit engagement is to gather sufficient appropriate evidence to determine whether the client's financial statements are relatively free of material misstatements due to error or fraud and to then provide an opinion on the truth and fairness of those financial statements in the auditor's report. Analytical procedures are a vital audit tool in achieving the objectives throughout the audit process and are extensively used in practice, according to several surveys (Hirst & Koonce, 1996; Tabor & Wills, 1985; Trompeter & Wright, 2010). The efficiency and effectiveness of analytical procedures allow auditors to attain competitive advantage and reduce the risk of legal liability. However, the audit environment is dramatically changing and it is becoming more challenging for auditors to apply analytical procedures efficiently and effectively.

Audit clients now embrace highly automated enterprise processes and are striving to employ advanced system and analytics tools in order to remain competitive and relevant in today's business environment. Frequently, their systems are integrated with the "cloud", the "internet of things", and social media, which produce huge volumes of disparate data. They are applying complex business analytical approaches involving "big data" and artificial intelligence so as to generate useful information for decision-making. The ability to discover variances, understand aggregate content, and to predict trends has clearly created a profound imbalance between clients and their auditors, since auditors have not turned their attention to advanced analytics tools to the same extent as their clients (Appelbaum, Kogan, & Vasarhelyi, 2017). Different industries, complicated businesses and international locations often obscure the ability of auditors to exercise professional judgment. Large amounts of data make it even more difficult for them to integrate all cues during the audit process.

Audit data analytics (ADA) has emerged from these rapidly evolving and varied corporate systems (Appelbaum et al., 2017). The scenario creates an opportunity and the urgency for auditors to move towards ADA and undertake more advanced predictive- and prescriptive-oriented analytics. In such a data-driven environment, a more advanced audit approach, reflecting the confluence of automation and advanced analytics, would directly improve the efficiency and effectiveness of analytical procedures through analytically reviewing and automatically identifying outliers for the subsequent audit examination in detail.

On the other hand, the issue of the expectation gap between investors and auditors has been an issue for a long time and could get worse if auditors do not give sufficient attention to ADA in the near future. Investors expect auditors to find frauds because they are concerned with the quality of financial statements on which a better decision is based. However, periodic high-profile cases of fraudulent financial reporting raise serious concerns over the credibility of regulators and auditors. In response, regulators recently have turned to ADA, which has opened up the possibility of fraud detection in a manner not previously possible, and of developing sufficiently precise expectations to identify a material misstatement in the context of analytical procedures (ICAEW, 2016).

Regulators have issued auditing standards to clarify auditors' responsibilities for detecting fraud prior to the issuance of a company's financial statements, and to require them to maintain an appropriate level of professional scepticism. However, the assessment of fraud risk and identification of material misstatements due to error or fraud in the financial statements could be a particularly challenging task to perform because frauds involve intentional deception and auditors may have little experience with actual instances of fraud. The scenario puts significant pressures on auditors. One response to these pressures is to enhance the efficiency and effectiveness of analytical procedures as an audit tool, with the highlighted demand for more precise expectation models. Because the standards do not prescribe a list of analytical procedure techniques for auditors in all audits, they have discretion to make their own decisions on whether traditional techniques, such as ratio analysis, or more complex analytics (e.g., ADA) should be utilised. Hence, it seems necessary to investigate the relative efficacy of these techniques and seek the most effective method, which possesses the best predictive ability and error-detection performance.

Clearly, manual and simple analytical procedure techniques are likely to be inefficient and ineffective in a data-rich context. Glover, Prawitt and Drake (2014) question whether auditors can efficiently and effectively conduct an engagement by using trend and ratio analysis and scanning, if audit clients are applying more advanced business analytic techniques. Similarly, Appelbaum et al. (2017) suggest that auditors should “expand their use of analytical procedures beyond that of scanning, ratio and time series analysis, and detailed examination” (p. 8). Further, the need for precise expectation models also calls for the heightened use of advanced analytical procedure methods. These more advanced approaches enable auditors to create accurate expectation models and can be used to generate predicted values with an allowable variance from actual accounting values for attention-directing purposes. The expectation models developed through the application of more advanced approaches need to be examined in greater depth.

Alles (2015) suggests that any adoption by auditors of advanced analytic techniques would be due to forces exogenous to the firm. Factors such as the increasing complexity of client transactions, analytics and data sources, and the subsequent increase of audit risk if manual and simple analytical procedures are used, will drive the application of advanced analytics by auditors. They have led to the recent revival of interest in ADA by such exogenous forces (Appelbaum et al., 2017) and the need to investigate ways to integrate more advanced analytics in their engagements, since the overall importance of advanced analytics is hard to ignore (Alles, 2015). The general goal for the adoption of ADA by auditors is to gain greater efficiencies and effectiveness in the audit process. Historically, auditors have been using diverse descriptive approaches but are now about to embrace more predictive and prescriptive analytics. They need to acquire a broad knowledge of advanced analytical techniques and get more comfortable with them. Consistent with this, more research is needed on advanced approaches, such as modern statistical and machine-learning methods, to assess their ability to develop precise expectations, efficiency and effectiveness for auditors in performing analytical procedures and deriving results in many situations.

Neural networks, a type of machine-learning method, have been said to possess superior pattern recognition and predictive abilities (Ngai, Hu, Wong, Chen, & Sun, 2011; West & Bhattacharya, 2016). Their power stems from their ability to model complex data relationships without underlying assumptions and deal with “noisy” data commonly found in practice (Hecht-Nielsen, 1989). The application of neural networks in business, including in the auditing domain, has been growing rigorously. A paper by Coakley and Brown (1993) was the first to consider and research the potential application of neural networks as an audit analytical procedure technique. They assessed the effectiveness of the neural network as an analytical procedure by comparing the performance of its counter-methods, such as financial ratios and regressions. They found that the neural network performed slightly more effectively than the other approaches, based on the sum of Type I and Type II error rates, and none of the three approaches provided a significant improvement in effectiveness over a purely random procedure. Whereas Coakley and Brown (1993) used case data from a single firm, the current study extends their work by using data from multiple firms listed on the Hong Kong Exchange in the China dairy industry. It also incorporates the effects of exogenous variables.

In summary, this study aims to evaluate the effectiveness of neural networks as an analytical procedure in a scenario involving multiple case firms. It achieves this objective by training and validating the models with data derived from six China dairy companies and also by testing the developed models with data obtained from one holdout China dairy company, the accounts of which have known issues. It contributes to theory by supplementing evidence for the conclusion of Coakley and Brown (1993). It also contributes to practice by demonstrating the potential use of neural networks and ADA in effectively performing analytical procedures via a practical case study.

The remainder of the thesis is as follows. The second chapter provides a review of the literature that has examined the effectiveness of analytical procedures and the application of neural networks in the auditing domain. The third chapter develops the study’s hypotheses. The fourth chapter describes the adoption of a methodology which follows Coakley and Brown’s (1993) as closely as possible, in order to render a fair comparison between the studies. The fifth chapter presents results of the study. The sixth chapter discusses the findings and limitations of the study. The final chapter contains some conclusions and recommendations for future research.

2 LITERATURE REVIEW

2.1 Analytical procedures

2.1.1 Use of analytical procedures

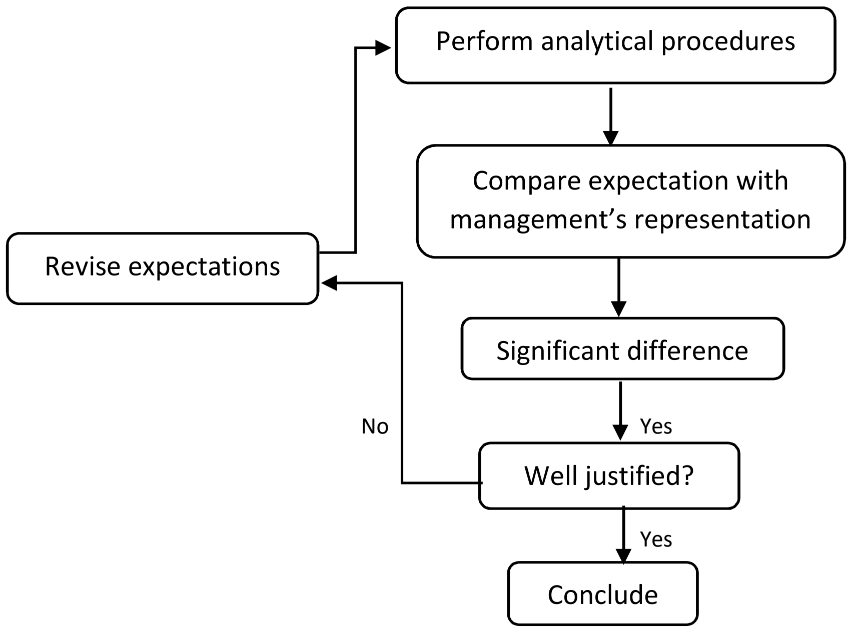
Statement on Auditing Standards (SAS) No. 56 *Analytical Procedures* in the United States (U.S.) is generally considered to be an authoritative pronouncement on analytical procedures in the literature (AICPA, 1988). SAS 56 defines analytical procedures as the "evaluation of financial information made by a study of plausible relationships among both financial and nonfinancial data" (AICPA, 1988, 56, para. 2) and that "analytical procedures involve comparisons of recorded amounts, or ratios developed from recorded amounts, to expectations developed by the auditor" (AICPA, 1988, 56, para. 5). International Standards of Auditing (ISA) No. 520 *Analytical Procedures*, issued by the International Auditing and Assurance Standards Board (IAASB), the equivalent of SAS 56, states that analytical procedures involve the evaluation of "financial information through analysis of plausible relationships among both financial and nonfinancial data" and necessarily investigating "identified fluctuations or relationships that are inconsistent with other relevant information or that differ from expected values by a significant amount" (ISA 520, paras. 4–5).

SAS 56 requires the application of analytical procedures in the planning and overall review stages of all audits and provides guidance on their use. SAS 99 *Consideration of fraud in financial statement audit* further explains the use of analytical procedures as risk-assessment procedures during the audit planning phase. It requires auditors to obtain an understanding of the entity and its internal control environment in order to assess and identify the risk of material misstatements due to error or fraud at the financial statements and assertion levels. Their equivalents, ISA 520 and ISA 315, require the same. In both U.S. and international standards, auditors are encouraged to use analytical procedures as an audit approach or a supplement to other approaches during the substantive testing stage, while being compulsorily required to apply analytical procedures in both planning and review stages, with the aim of assessing the risk of material misstatements and the consistency between the financial statements and the auditor's understanding of the entity, before drawing a conclusion on the financial statements.

The purpose, degree of precision and reliance on analytical procedures vary throughout the audit process. During the engagement stage, analytical procedures will be performed to understand the business and its environment and the transactions and events that have occurred since the last financial year. During the planning stage, analytical procedures will act as an attention-directing device to assess the risk of material misstatement due to error or fraud and identify the unusual trends and relationships, in order to subsequently design the nature, timing and extent of substantive audit procedures. During the substantive testing phase, analytical procedures may be applied on their own or to supplement tests of details to gather evidence about certain audit assertions relating to account balances or classes of transactions, when it is more effective or efficient than other tests of details in achieving a specific objective of the audit procedure. During the final review stage, analytical procedures will be employed to comprehensively assess the reasonableness of the financial statements, to ensure that no further irregularities or unidentified material misstatements exist and that an appropriate audit opinion is given. Cho and Lew (2000) argue that "attention directing", "test reducing", and "assessing fairness" are the three primary roles of analytical procedures.

A comprehensive model to clarify the process of the application of analytical procedures was developed by Hirst and Koonce (1996) and includes five stages: expectation development, explanation generation, information search and explanation evaluation, decision-making, and documentation. Bell, Peecher, and Solomon (2005) further illustrated the process in detail through a flowchart (Figure 1). All in all, a basic premise underlying the application of analytical procedures is that plausible relationships among data may reasonably be expected to exist and continue in the absence of conditions to the contrary (AICPA, 1988). Significant differences between the expected amount and recorded amount "are investigated in the hope of locating” (Colbert, 1994, p. 1) material misstatements due to error or fraud in the financial statements.

Figure 1 The process of risk assessment in relation to analytical procedures



Source: Bell et al. (2005, p. 23)

The importance of analytical procedures in the audit process is widely recognised. Hylas and Ashton (1982) demonstrated that analytical procedures are particularly useful as they can detect a significant proportion of material errors at the early stages of an audit engagement. Over time, an overall increase in the use of analytical procedures in general has been reported by numerous studies in various countries (Cho & Lew, 2000; Hirst & Koonce, 1996; Lin & Fraser, 2003; Smith, Psaros, & Holmes, 1999; Tabor & Wills, 1985; Trompeter & Wright, 2010). Reasons given for the increase include auditors being able to form more precise expectations and make better decisions due to the adaption of a business-risk audit methodology, technological advancements and the incorporation of non-financial information in an audit (Lin & Fraser, 2003; Trompeter & Wright, 2010).

Analytical procedures need to be efficient and effective if they are to help public accounting firms simultaneously reduce effort, costs and the probability of litigation in the context of a competitive environment. However, Green and Calderon (1994) considered that the 1987 Treadway Commission report, calling for changes in Generally Accepted Auditing Standards (GAAS) so as to “make greater use of analytical review procedures to identify areas with a high risk of fraudulent financial reporting” (p. 8), had resulted in a shift in focus in applying analytical procedures from audit efficiency to audit effectiveness. The effectiveness of analytical procedures is reflected in a superior error detection performance and predictive ability, which are the basis of properly directing auditors’ attention to possible errors and risk areas that require more thorough tests.

2.1.2 Error detection ability of analytical procedures

The error detection ability evaluates if an analytical procedure method can efficiently assess the absence of material errors and reliably identify the presence of material errors in financial accounts of the financial statements. Many prior studies have emphasised the significance of analytical procedures in detecting financial statements errors (Biggs & Wild, 1984; Hylas & Ashton, 1982; Tabor & Willis, 1985). A relationship between the error-signalling ability of an analytical procedure with the noise in recorded account balances was noted by Kinney (1987). He found that variations between expected and recorded values across several related financial ratios were useful in identifying the cause of an error. Therefore, the pattern of fluctuations across financial accounts may indicate the presence of material errors and signal the possible causes of the error. Given a level of noise, factors such as error dispersion patterns, error size to materiality level, confidence interval and investigation rule, have been shown to affect the overall error-signalling ability of analytical procedures (Kinney & Salamon, 1982; Knechel, 1986).

Regarding how to seed a material error in a simulated experiment, there is no clear way to assess whether an error is concentrated in one specified accounting period or spread over many periods. However, Wheeler and Pany (1990) examined the signalling performance by seeding a quarterly or annual material error amount. They concluded that the analytical procedures do signal very well when an annual material error was seeded into an individual quarter’s data.

The materiality level interactively relates to the size of the error detected during the audit process. The materiality threshold, the acceptable amount of potential misstatement, must be sufficiently small to enable the auditor to identify misstatements that could be material either individually or when aggregated with other misstatements, but also cannot be so small as to trigger excessive false positives. Empirically, there are different approaches to measuring materiality levels, such as measuring 10% of average earnings over a three-year period (Kinney, 1979), 10% of net income (Holstrum & Messier, 1982), 1.6 times (greater of total assets or revenue)$^{2/3}$ (Elliot, 1983), and 0.038657 multiplied by (revenue)$^{2/3}$ (Coakley & Brown, 1993, p. 6; Warren & Elliott, 1986). The determination of materiality level is a matter of professional judgment and is left to the auditor, in accordance with auditing standards.

The confidence interval represents the level of assurance desired from the procedures and is a range of likely values for the population parameters, based on the auditor's desired level of confidence. It is calculated based on alpha risk level, which is the risk of the incorrect rejection of the book amount. Alpha risk level of 0.33 is the level most frequently used in the literature.

An investigation rule is required to signal the need for further investigation of a significant difference between actual and expected values. The amount of difference falling within the range of minimum bound to maximum bound is acceptable without further investigation. Significant unexpected differences indicate a likelihood of material misstatement and should be investigated in detail and corroborated with management's explanations. Investigation rules in the auditing literature include percentage change rules and statistical rules. Kinney and Salamon (1982) formulated a rule based on upper confidence limits of the forecast errors. Kinney (1987) proposed a statistical rule to evaluate the effect of errors on financial ratios. Stringer and Stewart (1986) developed a statistical technique for analytical review (STAR) rule for the use of regression models, based on upper confidence limits adjusted for the anticipated dispersion pattern of accounting errors. The statistical investigation rule signals if the standardised difference between the recorded and the predicted account balance exceeds the critical Z-value that is based on the alpha risk level specified by the auditor. Generally, statistical rules using an assumed normal distribution are able to generate more precise investigation signals to assess whether material misstatement is likely at a specified level of assurance (Kinney, 1987; Wheeler & Pany, 1990).

From the auditor's traditional perspective, there are two types of decision risk, Type I and Type II error. The performance of analytical procedures is measured by the observed Type I error rates and Type II error rates. Type I error (an incorrect rejection) occurs when the method signals investigation of an account balance that is in fact correctly stated. Type II error (an incorrect acceptance) occurs when the method does not signal investigation of an account balance that actually contains a material error. The auditor desires a low Type I error rate (efficiency) and a low Type II error rate (reliability) because of unnecessary additional audit costs and an inappropriate audit opinion on the financial statements, respectively. Auditors are probably more concerned with reliability since Type II errors cause potentially higher costs due to litigation (Colbert & Wilson, 1991; Ngai et al., 2011). Loebbecke and Steinbart (1987) compared the sum of Type I and Type II error rates to a benchmark of 1.0. If the sum of Type I and Type II error rates is greater than or equal to 1.0, an analytical procedure would be no more effective than a coin flip in deciding whether to investigate a given account (Wheeler & Pany, 1990). However, it ought to be noted that Wheeler and Pany (1990) were concerned that summing the Type I and Type II error rates could double-count errors, since the generation of two errors for a single prediction is not realistic. As auditor's risk preferences for the trade-off between Type I and Type II errors vary, adding the Type I and Type II rates implicitly assumes that the auditor equally weights either type of error (Knechel, 1986).

2.1.3 Predictive ability of analytical procedures

Predictive ability measures how close an expectation developed by an analytical procedure method is to the recorded value in the financial statements. An expectation can be a specific predicted number (e.g., dollar amount or ratio value), a percentage, a direction or an approximation. Regardless of the stage of the audit, analytical procedures involve a comparison of the recorded amount against expectations developed by the auditor (AICPA, 1988, para. 5). Glover, Prawitt, and Wilks (2005) contended that the quality of the expectation affects the effectiveness of an analytical procedure. McDaniel and Simmons (2007) stated that the precision of an expectation refers to the quality of the expectation and is a measure of the closeness of the developed expectation to the actual amount (AICPA, 1988, para. 4). The audit judgments would benefit from explicit quantification of precision, consistent with the statistical concept of confidence interval, prior to combining cues, to promote improved integration of evidence (Lin, Fraser, & Hatherly, 2000). The greater the degree of precision, the greater the likelihood that the difference is identified as a misstatement (ACCA, 2016).

Professional standards clearly show the importance of precision for the effectiveness of analytical procedures (AICPA, 1995, 1998, 2004). Blocher and Patterson (1996) stated that the development of a precise expectation was a prerequisite to properly performing analytical procedures in accordance with auditing standards; otherwise, the procedure was potentially biased by other irrelevant information. Audit research on the topic has evolved in two research directions which are: (1) what information should be included in analytical procedures, and (2) what expectation model should be used (Chen & Leitch, 1998). Four key factors affecting the precision of a given expectation are: the predictability of the account, the level of data disaggregation, the reliability of the data, and the type of method used to form the expectation (Blocher & Patterson, 1996).

McDaniel and Simmons (2007) investigated how auditors assess expectation precision and found that auditors consider both account predictability and level of detail when determining the precision of an expectation and the likelihood of misstatements. Regarding account predictability, Wheeler and Pany (1990) observed the lowest error rates for those accounts (such as interest and depreciation) where the primary substantive procedure would be re-computation and found that the assertion of SAS 56 (that income statement accounts are more predictable than balance sheet accounts) was inconsistent with their evidence. Disaggregated data can be the composition of a balance, based on time and the source of the underlying data elements, and may be preferable in analytical procedures because of the improvement of prediction accuracy and the effectiveness of analytical procedures (ACCA, 2016). The types of high-frequency, disaggregated data in the literature are daily observations (Kogan, Alles, Vasarhelyi, & Wu, 2010), monthly data (Chen & Leitch, 1998; Cogger, 1981; Dzeng, 1994; Kinney, 1987; Knechel, 1988), and quarterly data (Wheeler & Pany, 1990).

External data are generally considered more reliable than internally generated data as they are independent of the firm being audited. Both industry and economic factors are critical for developing expectations when using analytical procedures because they have impact on organisational activities and accounts and are not affected by an accounting error (Chen & Leitch, 1998). Although Libby and Luft (1993) called for the incorporation of industry-specific experience factors in designing research, they are seldom used within auditing research (Hoitash, Kogan & Vasarhelyi, 2006). Prior research has suggested that the inclusion of external industrial and economic data can improve the predictive ability of analytical procedures (Lev, 1980; Loebbecke & Steinbart, 1987; Neter, 1980; Wild, 1987). Additionally, the inclusion of account balances from peer companies as independent variables in the expectation model was examined by Hoitash et al. (2006), who concluded that it helped considerably in detecting errors, while not always improving the prediction accuracy.

In summary, analytical procedures perform effectively if correct investigation signals can be produced, based on an objective and precise expectation that relationships among data reasonably exist. A better-quality expectation can be formed for a predictable financial account with detailed level data extracted from reliable sources, through using appropriate techniques, which will be discussed in the following subsection.

2.1.4 Analytical procedure techniques

Another research direction has been to investigate techniques to improve auditors' abilities to quantify and incorporate precision into investigation decisions (Kinney, McDaniel, & Martin, 2005; McDaniel & Simmons, 2007). SAS 56 and ISA 520 do not prescribe the list of analytical procedure techniques for auditors to use to develop expectations. Instead, they state that analytical procedures range from simple comparisons to complex models (AICPA, SAS 56, para. 2; ISA 520, para. 4). Because of that, Colbert and Wilson (1991) described the type of analytical procedure techniques used to predict account balances and argued that they are limited only by the availability of reliable data and the creativity of the auditor. The aspect these techniques have in common is to estimate the expected value by modelling underlying relationships (McKee, 1989).

Green and Calderon (1994) stated that SAS 56 requires auditors to use both qualitative and quantitative analytical procedures. Qualitative approaches consider three factors in a particular business situation (Loebbecke & Willingham, 1988). They are: (1) conditions allowing an irregularity, (2) motivation for committing an irregularity, and (3) personal attitudes admitting an irregularity. Quantitative analytical procedures are tools which are applied to quantify deviations between the expected amount and the recorded account balance. Qualitative analytical procedures answer the question "Why?" and quantitative analytical procedures signal what should be examined (Green & Calderon, 1994). Tabor and Willis (1985) suggested that the use of analytical procedures has shifted toward the quantitative procedures, in particular.

Non-statistical methods, such as account changes and simple trend and ratio analysis, were traditionally used to perform analytical procedures. Chen and Leitch (1998) argued that non-statistical analytical procedures are not able to produce objective results, as they are based on limited information. Statistical models perform better because of the incorporation of both structural relationships among accounting numbers and relevant exogenous variables, independent of accounting errors (Chen &

Leitch, 1998). Knechel (1988), Wilson and Colbert (1989), and Wheeler and Pany (1990) found that statistical analytical procedures requiring more relevant information offered significantly more accurate expectations and performed better in error detection.

Blocher, Krull, Tashman, and Yates (2002) separated approaches required to develop the explicit expectation (time series and regression) from other approaches used to develop the implicit expectation (trend and ratio analysis), and suggested that the explicit expectation approaches were more informed, precise and reliable than the implicit approaches. Simplely, Fraser, Hatherly, and Lin (1997) categorised three types of analytical procedure techniques: non-quantitative (NQT) or judgmental, such as scanning; simple quantitative (SQT), such as trend, ratio and reasonableness tests; and advanced quantitative techniques (AQT), such as time series, regression, and neural networks.

Several studies in the extant literature have developed various analytical procedure techniques and evaluated their effectiveness in developing precise expectations. McKee (1989) developed the Martingale model to calculate the expected account balance or ratio using the weighted average change method. Others have assessed regression analysis and suggested that it may be more effective than traditional procedures because of the considerable improvement in predictions (Akresh & Wallace, 1982; Kinney, 1978; Wilson, 1991). Dzeng (1994) introduced Vector Autoregression (VAR) as a possible analytical procedure technique and found that VAR performed slightly better than multivariate regression models. Kinney (1978) and Wilson, Colbert, and Minyard (1991) tested the effectiveness of Autoregression Integrated Moving Average (ARIMA) and found that ARIMA produced the smallest mean absolute error as well as the smallest prediction bias, as compared to regressions. Duguan, Gentry, and Shriver (1985) proposed the X-11 time-series model to develop expectations by incorporating trends and seasonality in account balances. However, Wheeler and Pany (1990) decided to use the X-11 model because ARIMA was criticised by Arrington, Hillison, and Icerman (1983), and found that the X-11 model was not significantly superior to multivariate regression models. Chen and Leitch (1998) concluded that the prediction performance of the structural model was not better than that of multivariate stepwise regression models through extending prior studies, such as those of Wheeler and Pany (1990), Wilson and Colbert (1989), and Lorek, Wheeler, Icerman, and Fordham (1995), who tested several error detection models without structural relationships, and Wild (1987) who used a structural model for one company. Loebbecke and Steinbart (1987) and Pany and Wheeler (1992) concluded that while these methods may be good at spotting fluctuations resulting from the presence of material monetary errors in the account balances, they do not reliably indicate the absence of material monetary errors.

More recently, Appelbaum, Kogan, and Vasarhelyi (2016) examined more than 300 papers published since the mid-1950s that discussed analytics in at least one phase of the audit and categorised techniques into the following five groups. They were: (1) audit examinations, such as transaction tests, ratio analysis, re-performance, and CAATs; (2) unsupervised, such as clustering, text mining, and process mining; (3) supervised, such as support vector machine, neural networks, genetic algorithms, and C4.5; (4) regression, such as logistic, linear, time series, ARIMA, univariate, and multivariate; and (5) other statistics, such as Benford's Law, structural models, and Monte Carlo simulation.

Based on the conclusion of prior experimental studies, AQTs promise more precise expectations, better error detection performance, and greater audit effectiveness compared to SQTs. Given the incorporation of relevant financial data, operating data, and industry and economic factors, AQTs are more effective because they consistently test the underlying assumption that changes in operation and environment primarily cause the changes in the financial statements. Among those evaluated AQTs, stepwise regression models appear reasonably effective, being efficient and reliable in identifying the absence of expected relationships or the presence of unexpected relationships by developing a precise expectation. Therefore, the current study selects ratios, stepwise regressions, and neural networks as the focal techniques to be compared.

These techniques have their own advantages and disadvantages and are ranked from lowest to highest in order of their inherent precision. Financial ratios have been the most widely used analytical procedure techniques over the years (Colbert & Wilson, 1991). However, ratio procedures generally produce ambiguous interpretations because it is difficult to separate fluctuations caused by errors from other normal fluctuations. Hence, this analytical procedure method has become inefficient because of unnecessary investigations increasing the overall cost of an audit. Stepwise regression models, as parametric models, rely on different underlying assumptions. Neural networks, as non-parametric models, are able to recognise data patterns and discover complex data relationships without any assumptions, even if the data are noisy or distorted. That is a major reason why many studies conclude that neural networks are likely to outperform regression models. Regressions might face modelling issues in high-dimensional spaces, such as collinearity, redundancy, relevance, etc. In contrast, providing a large number of input parameters to a neural network does not pose a model structure problem because the neural network will learn to ignore unimportant data by assigning near-zero values to their weights (Hecht-Nielsen, 1989). Further, outputs generated by neural networks are not sensitive to minor variations in the input patterns, which is very important when performing analysis of financial statements, since variations in the account balances always occur (Coakley & Brown, 1993; Hecht-Nielsen, 1989).

2.2 Neural networks

"Data mining" is defined as a process of applying statistical, mathematical, artificial intelligence, and machine-learning techniques to extract useful implicit information and gain knowledge from a database (Turban, Aronson, Liang, & Sharda, 2007). Data mining has been used extensively in the literature of financial fraud detection. Phua, Lee, Smith, and Gayler (2010) pointed out that fraud detection has become one of the best-established applications of data mining in industry. Kou, Lu, Sirwongwattana, and Huang (2004) highlighted the key point that data mining can be employed to develop a new class of models to identify new frauds not before detected by human experts. Ngai et al. (2011) and West and Bhattacharya (2016) performed a comprehensive literature review for the periods of 1997-2008 and 2004-2014 respectively, and showed that neural networks outperformed other mining techniques in terms of accuracy.

Currently, neural networks have been proven useful in analysing different problems, such as bond rating (Surkan & Singleton, 1990), insurance fraud detection (Bermúdez, Pérez, Ayuso, Gómez, & Vázquez, 2008), credit card fraud detection (Fisher, 1999; Mulqueen, 1996), bankruptcy prediction (Altman, Marco, & Varetto, 1994; Odom & Sharda, 1990), and financial statement fraud detection (Bai, Yen, & Yang, 2008; Bose & Wang, 2007; Cecchini, Aytug, Koehler, & Pathak, 2010; Dong, Liao, Fang, Cheng, Chen, & Fan, 2014; Glancy & Yadav, 2011; Hoogs, Kiehl, Lacomb, & Senturk, 2007; Huang, 2013; Humpherys, Moffitt, Burns, Burgoon, & Felix , 2011; Kirkos, Spathis, & Manolopoulos, 2007; Ravisankar, Ravi, Rao, & Bose, 2011). However, the application of neural networks in the auditing literature is limited. Wong, Bodnovich, and Selvi (1997) and Wong and Selvi (1998) analysed published research from 1988 to 1996 and classified only one article into the auditing domain. Vellido, Lisboa, and Vaughan (1999) analysed prior studies from 1992 to 1998 and categorised six articles into the auditing area. The review conducted by Koskivaara (2003) showed that the main application areas in auditing were material errors (Coakley, 1995; Coakley & Brown, 1991, 1993), management fraud (Green & Choi, 1997; Fanning & Cogger, 1998; Feroz, Kwon, Pastena, & Park, 2000), going concern issues (Anandarajan & Anandarajan, 1999; Etheridge, Sriram, & Hsu, 2000; Hansen, McDonald, & Stice, 1992; Koh & Tan, 1999; Lenard, Alam, & Madey, 1995), financial distress problems (Fanning & Cogger, 1994), internal control risk assessments (Davis, Massey, & Lovell II, 1997; Ramamoorti, Bailey, & Traver, 1999) and audit fee forecasts (Curry & Peel, 1998).

Coakley and Brown’s 1993 study was the first in which neural networks were applied as an analytical procedure to direct auditors’ attention to financial accounts containing material misstatements due to error or fraud. They assessed the effectiveness of neural networks as an analytical procedure by examining the performance of a neural network process within the framework of a single case study, consisting of data for a single firm over a 48-month time period and varying three factors (the size of the error, the statistical level of confidence and sources of error). The research deliberately seeded errors in each month corresponding to two common errors: unrecorded purchase and fictitious sales. Based on the sum of Type I and Type II error rates, they concluded that neural networks performed slightly more effectively than counter-methods (financial ratios and regressions), and none of the three approaches provided a significant improvement in effectiveness over a purely random procedure. They also acknowledged that their conclusion concerning the overall effectiveness of neural networks as an analytical procedure was tentative until many variations recommended in their research are investigated.

Coakley and Brown (1993) recommended that their preliminary experiments should be extended to expand the dataset used for neural networks by including multiple firms and multiple financial periods for each firm. Also, they suggested that the neural networks could be more robust compared to other approaches when the input data were derived from audited periodic financial statements versus unaudited monthly account balances. Furthermore, they speculated that the regression method may reduce the explanatory ability for the sales and cost of sales accounts when external economic indicators were not considered. The current study was inspired by both their recommendations and the suggestions from prior research and extends the work of Coakley and Brown (1993) by incorporating data obtained from both audited periodic financial statements of multiple firms and from exogenous variables.

3 HYPOTHESES DEVELOPMENT

The purpose of this chapter is to develop the seven hypotheses to be examined in the current study. The evaluation of the effectiveness of neural networks as an analytical procedure was performed for a single case firm by Coakley and Brown (1993). The current study extends the prior study so as to assess the effectiveness of neural networks as an analytical procedure technique for multiple firms' financial data. Accordingly, it is hypothesised that the current study will have findings similar to those of Coakley and Brown (1993).

Coakley and Brown (1993) found that neural networks were slightly more effective than the financial ratio and regression approaches and that none of the three methods produced a significant improvement in effectiveness over a purely random procedure, based on the combined error rates. Before reaching these conclusions, they assessed the performance of the three methods by manipulating the effect of error sizes, statistical confidence levels and error sources for both inventory assumptions and by applying the developed models to the dataset. Their experimental results were as follows:

- When errors less than the materiality level were seeded, the financial ratio approach resulted in a lower Type I error rate, and the regression and neural network approaches resulted in higher Type I error rates.
- When larger errors were seeded, the three methods became more reliable.
- The three methods had a similar improvement in overall effectiveness across various error sizes.
- When the alpha risk was increased, the neural network approach was less sensitive as it produced more precise expectation than alternative approaches.
- The three methods had a similar improvement in overall effectiveness when varying the alpha risk.
- The regression and neural network methods seemed to have similar effectiveness with the purchases not recorded error, but a completely opposite effect with the fictitious sales error.
- The regression and neural network methods differently produced more effective signalling for certain accounts than the rest of the financial accounts.

- The three methods had a similar improvement in overall effectiveness across both error types with the different trade-offs between the Type I and Type II error rates.
- When applying the developed models to the dataset which was used to estimate the models, the neural network approach produced no Type I or Type II errors. However, the regression approach still produced higher average error rates, although its effectiveness was improved.

Consistent with the above, the following seven hypotheses were proposed in the current study based on both the conclusions and experimental results of Coakley and Brown (1993). All hypotheses are meant to be helpful in assessing the effectiveness of neural networks as an analytical procedure and are stated in alternative form.

H1: The neural network approach is more effective than the two alternative approaches.

H2: None of the three approaches provide an improvement in overall effectiveness over the benchmark of 1.0.

H3: The three approaches have increased error rates when the size of the immaterial error increases.

H4: The three approaches have reduced error rates when the size of the material error increases.

H5: The neural network approach has less trade-off between Type I and Type II errors than alternative approaches when varying the alpha risk.

H6: The regression and neural network approaches react similarly to the purchases not recorded (P) error, but not to the fictitious sales (S) error.

H7: The neural network approach generates investigation signals for all seven financial accounts to support assertions of a financial analyst.

4 METHODOLOGY

This study used an experimental design to examine the relative effectiveness of two well-documented analytical review techniques (ratio analysis and regression analysis) and an alternative approach, artificial neural networks. Archival data were obtained from six listed Chinese companies operating in the dairy industry in order to train and test alternative techniques. A seventh dairy company with a high likelihood of misstated financial statement accounts was also used as a novel holdout sample. The methodology for the study is discussed in detail in five subsections. First, basic information relating to case firms and the dataset is described. Next, sources of error are defined. The predetermined investigation rules are then discussed in the following subsection. The fourth subsection elaborates on the three analytical procedures used in the study, financial ratio, regression and neural networks. The final subsection provides details about the nature of the experiment, including the simulation used to evaluate the relative effectiveness of the three forms of analytical procedure.

The methodology adopted by Coakley and Brown (1993) was followed as closely as possible. All other design features were structured to parallel the prior study, except for the following specified differences. The main differences between two studies were: (1) the consideration of relevant exogenous factors; (2) unscaled outputs of neural networks; (3) basis of model comparison; and (4) ways to test the developed models. The prior study used solely financial data from both the balance sheet and income statement to develop regressions and neural networks with scaled output targets; compared regressions to neural networks based on adjusted R^2s; and then tested models using the existing dataset which was deployed previously to train and validate models. This study, (1) develops regressions and neural networks with financial, operational, industrial and economic data; (2) does not scale the output values of neural networks; (3) compares regression models and neural networks based on Akaike information criteria (AIC)[1]; and finally, (4) tests models using a dataset that is independent from that which was used to develop the models.

4.1 Description of case firms

In order to better generalise the models developed in this study, multiple Chinese firms were selected from a specific industry and stock exchange. Looking at a single industry can facilitate information transfer (Hoitash et al., 2006) and result in prediction improvements (Loebbecke & Steinbart, 1987; Wheeler & Pany, 1990). ChinaAg (http://www.chinaag.org/hkg-listed/) shows that 15 Chinese companies listed on the Hong Kong Exchange engage in production of dairy products. The list of case firms is shown in detail in Appendix I. Three dairy companies are not major players in China's dairy market and have been very inactive. Four companies do not have dairy business as their main operation. One company offers financial services relating to dairy. As a result, the remaining seven dairy companies were included in this study. The companies under study engage in the production of raw milk, liquid milk and milk powder. These companies are independent firms under different management, and trade dairy products (particularly raw milk) with each other in the relatively concentrated China dairy market. Their business relationships probably have a limited impact on data used to build models in this study.

Semi-annual data were used in this study due to their availability, although prior research suggests monthly data are preferable. Semi-annual data are considered reliable because reviews by independent auditors and semi-annual regulatory filings are expected to induce greater accuracy in a firm's half-yearly data than in its monthly data. Interim reports and annual reports of these companies were obtained for the period from 30 June 2013 through to 30 June 2017. No changes of accounting methods or other unusual events came to the researcher's attention in any of the sets of financial statements

[1] AIC is a type of fit statistic, which is used as a way of stopping criteria and model selection biases. http://documentation.sas.com/?docsetId=emhpprcref&docsetTarget=emhpprcref_hpreduce_details05.htm&docsetVersion=14.2&locale=en

over the period. Year-end audited financial statements were generally given a true and fair view opinion over the period, except one firm for the financial year ended 31 December 2013. However, its balances from 30 June 2013 onwards were not affected by the qualified audit opinion because only its opening balances (profit and cash flows for the year ended 31 December 2012) were qualified. Therefore, the half-yearly balances used in this study were reasonably assumed to be free of material misstatement and error.

The dataset concerns financial, operational, industrial and economic factors. Financial data relate to account balances in the balance sheet, income statement and cashflow statement. Operational data are the number of employees and cows. Industrial data include the quantity of dairy goods produced, milk yield rate, raw milk price index and alfalfa market price. Economic data contain the China inflation rate and 3-month interbank interest rate. Revenue ranges from RMB 385,129,000 to RMB 29,465,755,000 and property, plant and equipment ranges from RMB 283,640,000 to RMB 12,698,374,000 for case firms.

Five accounts directly related to the sources of errors were included and are discussed in the next subsection. Additional accounts included in the study were predictor variables identified by stepwise regression models. In total, 22 accounts and aggregates were selected for the study. Their means, medians, standard deviations and coefficients of variation for the period are shown below (Table 1). Figure 2 shows the change of the averaged account balances for these five accounts over the period. The movement of cost of sales closely followed the fluctuation of sales. All accounts showed a gradually upward growth trend. However, the growth rate of the inventory account declined since financial year 2016.

4.2 Definition of sources of error

To provide evidence of the effectiveness of various analytical procedure techniques used in auditing, experiments require sources of error to examine their error detection ability. The accounts of sales and cost of sales are often targeted for manipulation and fraud. Effects could be severe if frauds are committed through these accounts (ACFE, 2016). Thus, two sources of error used in the current study were unrecorded purchases and fictitious sales. They are among the sources of misstatement and error that occur most frequently in practice (Coakley & Loebbecke, 1985) and their affected accounts often require adjustments (Kinney, 1987).

Unrecorded purchase of merchandise on account and fictitious sales on account could occur due to cut-off error and deliberate misstatement. Cut-off error occurs when a legitimate transaction is recorded in the wrong accounting period. For instance, this is where merchandise purchased and delivered before the end of the accounting period is recorded after the end of that accounting period, or where merchandise sold and delivered after the end of the accounting period is recorded before the end of accounting period. Cut-off errors are self-correcting because the account for the accounting period will be understated or overstated and then the account for the next accounting period will be overstated or understated, respectively. Cut-off errors cancel each other out when the two accounting periods are combined. Deliberate overstatement and understatement are motivated by a particular incentive, which leads to fraudulent financial reporting.

Table 1 Financial accounts included in the study

Financial account	Mean	Median	Standard deviation	Coefficient of variation
Sales	5,267,758	1,383,627	9,019,854	1.71
Trade receivables	615,683	185,396	723,870	1.18
Cost of sales	3,520,515	746,162	6,215,010	1.77
Trade payables	1,252,935	337,479	1,844,036	1.47
Inventory	1,054,846	672,904	1,196,105	1.13
Inventory write-down	14,267	0	39,884	2.80
Selling expenses	1,115,700	145,627	2,079,828	1.86
Administrative expenses	233,425	84,678	355,353	1.52
R&D expenses	16,160	0	28,802	1.78
Fair value change on biological assets	-60,877	0	133,985	2.20
Cash and cash equivalents	2,003,218	870,134	2,605,464	1.30
Short-term loans	1,865,115	918,404	2,063,224	1.11
Prepayment	266,712	77,454	439,926	1.65
Property, plant and equipment	3,420,051	1,849,250	3,517,202	1.03
Accumulated depreciation	1,719,661	239,429	2,967,404	1.73
Pledged property, plant and equipment	150,092	52,300	190,618	1.27
Cashflow from operating activities	467,542	167,893	770,829	1.65
Cashflow from financing activities	555,235	179,284	1,651,976	2.98
Current assets	5,433,842	2,569,502	6,783,160	1.25
Current liabilities	4,517,465	2,028,302	5,402,112	1.20
Number of employees	9,170	3,713	13,400	1.46
China inflation rate	1.94	1.90	0.46	0.23

Companies generally perform a stock-take at the end of financial year and adjust the book value of inventory accordingly to match what is actually on hand. The effect of cut-off errors and misstatements on the financial accounts changes when the stock-take is performed and inventory account is subsequently adjusted to match the physical count. If the merchandise is physically present at the time when a stock-take is performed, it will be included in the adjusted inventory amount. The inventory amount will not be adjusted if the stock-take is not performed at financial year end. An analysis of the two sources of error was provided for each of the two inventory conditions (Table 2).

When purchases of merchandise are not recorded, both inventory and payables are understated by the amount of unrecorded purchases, which consequently lead to the understatement of aggregated accounts, such as current assets and current liabilities. However, if a stock-take is performed and the inventory is adjusted to match the physical count, the error becomes an understatement of cost of sales. In terms of aggregated accounts, only current liabilities are understated. When fictitious sales are recorded, both sales and receivables are overstated by the amount of fictitious sales. The cost of items sold associated with fictitious sales would be recorded to the inventory and cost of sales accounts. The error and misstatement in net income will generally equal the gross margin multiplied by the amount of fictitious sales. At an aggregation level, current assets are also affected. However, if a stock-take is performed and the inventory account is adjusted to match the physical account, the misstatement of inventory and cost of sales are corrected. Only sales and receivables are overstated.

Figure 2 Averages of financial account balances

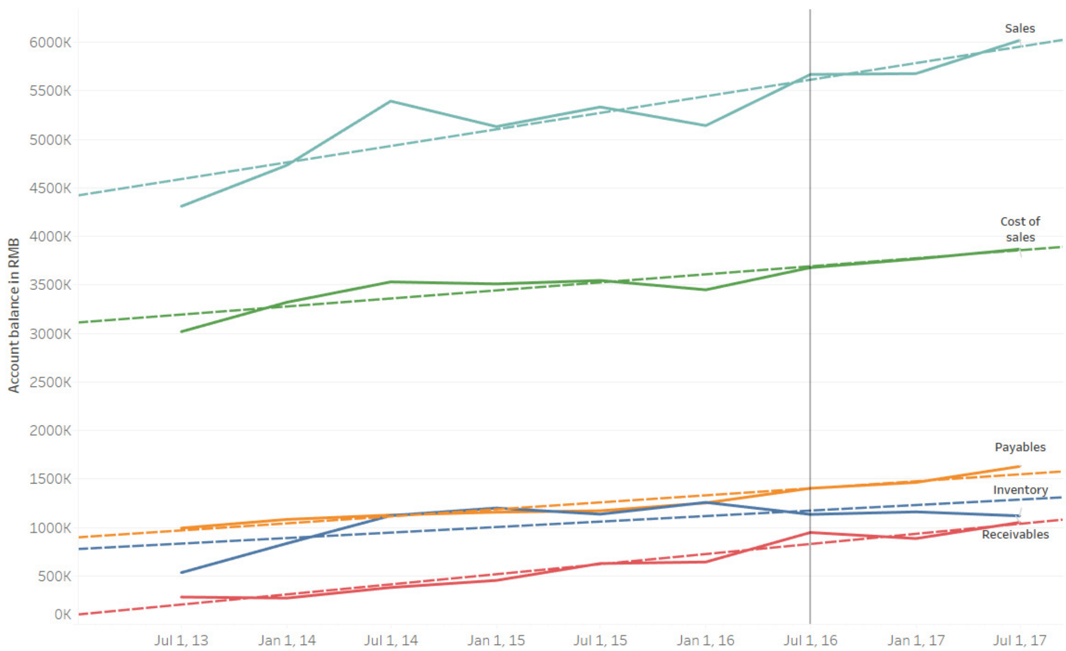


Table 2 Impact of source of error on financial accounts

Financial accounts	Source of error			
	Unrecorded purchase		Fictitious sales	
	Inventory adjusted to physical count	No physical count of inventory	Inventory adjusted to physical count	No physical count of inventory
Receivables	-	-	Overstated by S	Overstated by S
Inventory	-	Understated	-	Understated by C
Payables	Understated	Understated	-	-
Sales	-	-	Overstated by S	Overstated by S
Cost of sales	Understated	-	-	Overstated by C
Current assets	-	Understated	Overstated by S	Overstated by S-C
Current liabilities	Understated	Understated	-	-

Source: Coakley & Brown (1993), p. 5, S: Sales revenue for items sold, C: Cost associated with items sold.

4.3 Investigation rule

Financial statements of firms comprise the reported balances of various accounts that consist of many individual transactions over the period. An auditor performs analytical procedures in the stage of audit planning (ISA 315, para. 6b) and final review (ISA 520, para. 6) to assess the risks of material misstatement and error, identify the existence of unusual transactions and unexpected relationships, and ensure account balances are free of material misstatements and errors which otherwise need to be adjusted in the financial statements. The expected balance in an account at a particular time t (y_t) is estimated and compared with the actual reported balance (x_t) to determine if there is a significant discrepancy (e_t).

Equation 1: $e_t = x_t - y_t$

An auditor will take an action if the discrepancy exceeds the pre-determined materiality level (M) and confidence interval of the error (1-α). The materiality level used in this study is predefined based on the formula in Equation 2 developed by Warren and Elliot (1986) (Coakley & Brown, 1993, p. 6). Icerman and Hillison (1991) compared material amounts derived from that formula with empirical data collected from 49 manufacturing firms, and reported that the formula consistently produced material amounts which approximated the size of individual errors resulting in adjustments to the financial statements.

Coakley and Brown (1993) stated that materiality level derived from the formula in Equation 2 is reasonable if Icerman and Hillison (1991) findings are relied on.

Equation 2: Materiality (M) = $0.038657*(\text{Revenues})^{2/3}$

The alpha level (α) (also known as Type I error) controls the acceptable probability of concluding the account is in error when it is not in error, which is used to define the rejection zone (Chen & Leitch, 1998). The upper and lower bound limits of a 1-α confidence interval for the error ($UBL_{1-\alpha/2}$ and $LBL_{1-\alpha/2}$) are calculated. Further investigation will be required if the $UBL_{1-\alpha/2}$ exceeds M or the $LBL_{1-\alpha/2}$ is less than -M.

Coakley and Brown (1993) listed four possible decision outcomes for a given materiality threshold and confidence interval, as shown in Table 3. A Type I error is an incorrect decision to investigate if the analytical procedure signals the existence of a material error when none is present in the account. A Type II error is an incorrect decision not to investigate if the analytical procedure fails to signal the existence of a material error when the account is actually in error. The Type I error rate measures the efficiency of the audit execution since additional unnecessary accounts may be investigated. The Type II error rate measures the reliability of the analytical procedure and represents a detection risk in the audit risk model which is the risk of concluding that an account balance is free of material misstatements and errors when it actually is not. If an effective analytical procedure is used to signal investigations, the expected value of the sum of the Type I and Type II error rates would be less than 1.0, which is the sum of error rate when a purely random process is used (Loebbecke & Steinbart, 1987). Therefore, the effectiveness of the analytical procedure can be assessed by comparing the sum of the Type I and Type II error rates to a benchmark value of 1.0.

Overall, a lower Type I error rate would indicate that the analytical procedure is more efficient. A lower Type II error rate would indicate that the analytical procedure is more reliable. A lower sum of Type I and Type II error rates would indicate that the analytical procedure is more effective. The benchmark value of 1.0 is used to determine the effectiveness of analytical procedures, which represents the sum of the Type I and Type II error rates when a purely random process is used to signal investigations.

4.4 Analytical procedure methods

Financial ratios, regression and neural networks were used as the analytical procedures in this study. Their results were compared to assess the relative performance in terms of efficiency, reliability and effectiveness.

Table 3 Types of attention-directing decisions

	Results of analytical procedure	
Size of error in account balance	Investigation not signalled	Investigation signalled
Less than materiality threshold	Correct decision	Type I error
Greater than materiality threshold	Type II error	Correct decision

Source: Coakley & Brown (1993), p. 3.

4.4.1 Financial ratio procedure

Receivables turnover, inventory turnover, cost of sales ratio, accruals ratio and quick ratio were used, as shown in Table 4. Prior studies have evaluated the effectiveness of these ratios as analytical procedures and have demonstrated the ability of these ratios to reflect changes in the account balance due to the sources of error being investigated in this study (Coakley, 1982; Coakley & Brown, 1993; Kinney, 1987). According to prior research, financial ratios may be useful in directing an auditor's attention to material errors, although they cannot be completely reliable. Ending balances were used to compute five ratios, instead of average balances. If a material error sits in the ending balance of an account, averaging the beginning balance with the ending balance would make it more difficult to detect errors due to the reduced effect.

A statistical rule similar to that proposed by Kinney (1987) was used to evaluate the effect of errors on financial ratios. Let r_t and $\hat{r}_t$ represent the book value of a ratio and the expected value of the ratio at a particular time t respectively. $\hat{r}_t$ is assumed to be the average audited value of the same ratio for the previous audit year[2]. Let s_r represent the standard deviation of the audited values of ratio r for the prior audit year. If the standardised difference between book values (r_t) and expected values ($\hat{r}_t$) at a particular time t is so material and unlikely to be created by chance, the decision to investigate would be warranted.

Equation 3: Test statistic = $(r_t - \hat{r}_t)/s_r$

If the distribution of standardised changes is normal, a Z-value based on an alpha risk level specified may be used as a pre-set critical value for the deviation ($Z_{1-\alpha/2}$ and $Z_{1-\alpha/2}$). An auditor should investigate the accounts comprising the ratio if the calculated test statistic in Equation 3 is greater than $Z_{1-\alpha/2}$ or is less than $-Z_{1-\alpha/2}$, where α is the probability of a Type I error.

4.4.2 Regression procedure

The actual observed value of y is computed in Equation 4 and varies about the true mean value with variance σ^2. The parameters for the model, β, are derived from the validating dataset and then applied to predicted values in Equation 5. A predicted value of an individual observation will be represented by $\hat{y}$ in Equation 5. The difference between the observed value y and the predicted value $\hat{y}$ in Equation 6 can be used to establish a prediction interval. An auditor can be at least $1-\alpha/2$ confident that the misstatement and error is less than a material amount (M) if the upper bound limit of the prediction interval is less than M or/and the lower bound limit of the prediction interval is greater than -M.

[2] Kinney (1987) and Coakley and Brown (1993) calculated $\hat{r}_t$ based on the average value of last year's ratio adjusted for changes in the industry.

Table 4 Impact of source of error on evaluated financial ratios

Financial ratios	Source of error: Unrecorded purchase – Inventory adjusted to physical count	Unrecorded purchase – No physical count of inventory	Fictitious sales – Inventory adjusted to physical count	Fictitious sales – No physical count of inventory
Sales/Receivables	-	-	Up	Up
Cost of sales/Inventory	Down	Up	-	Down
Cost of sales/Sales	Down	-	Down	Up
Receivables/Payables	Up	Up	Up	Up
(Cash+Receivables)/Current liabilities	Up	Up	Up	Up

Source: Coakley & Brown (1993), p. 8.

Equation 4: $y = x\beta + \xi$

Equation 5: $\hat{y} = x\beta$

Equation 6: $Var(y- \hat{y}) = (1+x_0(x'x)\text{-}1x_0)\ \sigma^2$

Regression models were developed for five accounts (sales, cost of sales, inventory, receivables and payables) and two aggregates (current assets and current liabilities), on which the sources of errors have a direct impact. Stepwise regression techniques were used with current and lagged one and two periods of the financial data and a period indicator as the independent variable. The resulting models, associated multiple coefficient of determination (adjusted R^2) and AIC are shown in Table 5. The regression model provides a reasonable explanation for each targeted account and aggregate. Surprisingly, selected industrial indicators were excluded by regression models, probably because China dairy companies were severely affected by the 2008 Chinese milk scandal and their product prices were temporarily influenced by the international dairy product price, such as Fonterra WMP (DBS, 2017).

4.4.3 Neural network procedure

How neural networks should be designed depends on the problem demanding a solution and trial-and-error experiments. This study attempted to use neural networks to develop an expectation model for analytical procedures. To better design neural networks, it is necessary to carefully consider architectures, learning algorithms and training processes. Finally, the statistical decision rules are defined for the neural network to perform an analytical procedure.

4.4.3.1 The neural network architecture

The architecture of the neural network defines how nodes in a series of layers of a network connect to each other. There are different neural network architectures, such as multilayer perceptron, radial basis function networks, recurrent networks and others. This study used a multilayer perceptron (MLP) which has been used by most of the prior research in accounting and auditing (Coakley & Brown, 2000). With the MLP architecture, processing each node in a layer moves in the forward direction to every node in the next layer. When dealing with prediction issues, Coakley and Brown (2000) suggested that the number of input neurons "corresponds to the number of lagged observations used to discover the underlying time-series patterns" (p. 133) to capture the autocorrelation structures in the data. A systematic method, genetic algorithms, has been used in a number of studies to determine the optimal number of lagged input variables to include in the neural network (Zhang, Patuwo, & Hu, 1998). The number of output nodes to use depends on the number of expected values needed in the performance of analytical procedures.

The hidden neurons play a very important role in extracting the most useful features from the input vector and predicting values on the output vector (Ramamoorti et al., 1999). The number of hidden layers and hidden neurons directly affect the overall performance of a neural network. A neural network without hidden layers, similar to a generalised linear model (GLM), will roughly offer linear prediction. Generally, a neural network with a single hidden layer is sufficient to approximate a function that may extract the meaningful information from the data and attain the desired accuracy (Hecht-Nielsen, 1989). A neural network needs at most two hidden layers in order to approximate a particular function to a desired accuracy (Cybenko, 1988; Lapedes & Farber, 1987). Reed (1993) stated that the smallest neural network with an acceptable performance would be preferred because it has good generalisation and less likelihood of overfitting.

There is no rule of thumb in deciding the number of hidden neurons in each hidden layer. If too few neurons are used, the neural network will be unable to model complex relationships from data points. If too many neurons are used, the neural network may run a risk of overfitting. In the literature of accounting and auditing, researchers have adopted different heuristics or algorithms to determine the number of hidden neurons. Diamantaras and Kung (1996) argued that the number of hidden neurons in each hidden layer should not exceed the number of input variables in order to perform data compression and prevent the neural network from memorising the input data. Salchenberger, Cinar, and Lash (1992) recommended that the number of hidden neurons in a hidden layer should be 75% of the number of input variables. Subramanian, Hung, and Hu (1993) suggested that the number of hidden neurons in each hidden layer should fall within the range from k to n + 1, where k is the number of output variables and n is the number of input variables. Masters and Schwartz (1994) proposed that the number of hidden neurons in a single hidden layer equals sqrt (n*m), where n is the number of input neurons and m is the number of output neurons. If two hidden layers are used, the number of hidden nodes in the first hidden layer is mr^2 and the number of hidden nodes in the second layer is mr, where the parameter $r = (n/m)^{1/3}$. Cybenko (1989) and Hecht-Nielsen (1989) argued that a neural network would require at most (2n + 1) hidden neurons in each hidden layer to achieve the desired accuracy, where n is the number of input variables. Another consideration is that the increase in the number of hidden neurons in each hidden layer would lead to an increase in the number of connection weights, which eventually might be more than the number of training samples and constrain the network (Coakley & Brown, 2000).

Table 5 Parameters and coefficient of regression models[3]

Financial account	Model	Adjusted R^2	AIC
$Receivables_t$	=-22142	0.7626	1106
	+ 0.1874 Property, plant and $equipment_t$		
$Inventory_t$	=294164	0.8807	1115
	+ 81.5867 $Employees_t$		
$Payables_t$	=-27974	0.9920	1044
	+ 0.3722 Accumulated $depreciation_t$		
	+ 0.0963 Cash and cash $equivalents_t$		
	+ 0.0904 Short-term $loans_t$		
	+ 2.0071 Trade receivables $provision_t$		
	+ 0.3593 Cashflow from operating $activities_t$		
	+ 0.2971 $Prepayment_t$		
$Sales_t$	=-793798	0.9940	1164
	+ 570.5 $Employee_t$		
	+ 54.2093 R&D $expenses_t$		
Cost of $sales_t$	=-239661	0.9986	1076
	+ 0.1396 Property, plant and $equipment_t$		
	+ 413 $Employees_t$		
	+ 18.5671 R&D $expenses_t$		
	+ 3.0359 Trade $receivables_{t-1}$		
	- 0.6144 Cashflow from operating $activities_t$		
	- 283.4 $Employee_{t-1}$		
	+ 0.6221 Cost of $sales_{t-1}$		
	- 0.4061 $Inventory_{t-2}$		
	+ 0.2489 Trade $payables_{t-2}$		
Current $assets_t$	= 3592701	0.9984	1088
	+ 0.8111 Cash and cash $equivalents_t$		
	+ 0.2399 Short-term $loans_t$		
	- 1066686 China inflation $rate_t$		
	+ 1.4755 Fair value changes of $bio\text{-}assets_t$		
	+ 1.5002 Selling $expenses_t$		
	- 3.2468 Inventory $write\text{-}down_t$		
	+ 1.1121 $Inventory_{t-1}$		
	- 1.0329 Accumulated $depreciation_{t-2}$		
	+ 0.2356 Short-term $loans_{t-2}$		
	+ 0.1408 Cashflow from financing $activites_{t-2}$		
	- 725955 China inflation $rate_{t-2}$		
	+ 186.6 $Employee_{t-2}$		
Current $liabilities_t$	=93452	0.9981	1075
	+ 1.1253 Short-term $loans_t$		
	+ 123.7 $Employee_t$		
	-2.2232 Fair value changes of $bio\text{-}assets_t$		
	+ 0.8878 Selling $expenses_t$		
	+ 0.6490 Pledged property, plant and $equipment_{t-1}$		
	- 88.7791 $Employee_{t-1}$		
	+ 1.9064 Administrative $expenses_{t-1}$		
	+ 0.6983 $Prepayment_{t-1}$		
	+ 0.5816 $Prepayment_{t-2}$		

[3] Supplementary analysis was done with only lagged variables, which resulted in the lower R^2 and the higher AIC values. When applying AIC based model comparison, the regression approach would underperform the neural network method. Had these regressions been used, the performance of regression models would be different in detecting both immaterial and material errors. It might have a lower Type II error rate, but a higher Type I error rate.

Except for the aforementioned heuristic approaches, researchers in the literature have developed algorithms to adaptively determine the appropriate neural network. Some of algorithms start with a small number of hidden neurons and connections, then adaptively increase them until the desired result is obtained. For example, Fahlman and Lebiere (1990), using the Cascade-Correlation (Cascor), began with no hidden neurons and then incrementally added hidden neurons to improve the network's learning ability. Fanning and Cogger (1994) developed the Generalized Adaptive Neural Network Architecture (GANNA) where the architecture of the network evolutionarily grew with new hidden neurons to improve performance. Others started with a large network and pruned the number of connections during the training. For instance, Armstrong, Dwelly, Liang, Lin, and Reynolds (1991) developed the Adaptive Logic Network (ALN), which is a Boolean logic-based tree framework, evolved through deleting those branches that were not providing any additional information. Jhee and Lee (1993) employed a weight decay technique which prevents the network from growing overly large. It can be seen as a gradient descent on a quadratic regularisation term to penalise network complexity (Goodfellow, Bengio, Courville, & Bengio, 2016).

For the current study, the MLP architecture was employed. Stepwise regression models retained 22 input variables that had a high degree of predictive power while discarding others that did not. Neural networks were trained with 44 input nodes which represented the 22 input variables for the prior two lagged periods, and time period variable. Seven nodes in the output layer represented sales, cost of sales, inventory, receivables, payables, current assets, and current liabilities. Different experiments were conducted for neural networks to configure with hidden layers (0, 1, 2) and hidden neurons (11, 22, 45). Weight decay was applied to add a regularisation term to the cost function.

4.4.3.2 Selection of the learning algorithm

Machine-learning algorithms can be broadly divided into unsupervised and supervised, according to the kind of experience they are allowed to have during the learning process (Goodfellow et al., 2016).

Supervised learning algorithms attempt to associate some inputs with some outputs, given a training set of examples of inputs and outputs. The most frequently used supervised learning algorithm in the accounting and auditing domain is backpropagation with a gradient descent (Coakley & Brown, 2000). An optimal set of weights minimising a loss function represents knowledge within the neural network (Coakley & Brown, 2000). The error backpropagation process comprises two passes through the different layers of the network: a forward pass and a backward pass (Ramamoorti et al., 1999). The forward pass is to apply an input feature vector to the nodes of the network and propagate its effect through the network with a fixed set of weights. On the other hand, the backward pass is to propagate backward the error (the response of the network is subtracted from a desired response) through the network from the output layer, against the direction of gradients, to update the weights proportionally till the loss function is at the minimum.

When applying the gradient descent, backpropagation algorithm, the apparent disadvantages are the slow training process and the possibility of a local minima. However, the slow training process may not be a concern anymore because computers have significantly increased processing power relative to a decade ago. To avoid the local minima in reaching the global minimum, the use of optimisation algorithms (e.g., stochastic gradient descent, SGD) and appropriate adjustment of learning rate and momentum rate can be very helpful. SGD is an extension of the gradient descent algorithm and can approximately estimate the gradient by using a small set of samples (Goodfellow et al., 2016). The optimisation algorithm may not be guaranteed to arrive at a local minimum in a reasonable amount of time, but it often finds a very low value of the cost function quickly enough to be useful (Goodfellow et al., 2016). Choosing a proper learning rate can be difficult. Low learning rate will increase training time to converge because of tiny steps towards the minimum of the loss function. On the other hand, high learning rate will lead to oscillation without convergence because the neural network may overshoot to miss the minimum of the loss function or get stuck within a large dent or ravine of the surface (Coakley

& Brown, 2000). Momentum rate enables the speeding up of training but with a reduced risk of oscillating due to its relationship with learning rate indicated in Equation 7, where η learning rate and α momentum rate are used for the n-th correction for weight w_{ij} (Moreira & Fiesler, 1995, p. 2).

Equation 7: $\Delta\omega_{ij}(n) = -\eta[\partial E(n)/\partial w_{ij}(n)] + \alpha\Delta\omega_{ij}(n-1)$

The choice of an error function ultimately depends on whether the problem to be tackled is classification or prediction, on which model fit statistics and performance assessment are based. Coakley and Brown (2000) mentioned that the sum of square error (SSE) error function is most widely applied in the literature. However, Vellido et al. (1999) noted that the SSE error function has been used by papers under their review irrespective of them dealing with a classification or prediction problem. The SSE error function is suitable for prediction problems. The cross-entropy and softmax error functions are better for classification issues (Bishop, 1995; Bridle, 1990).

The transfer function is used to derive the output of a neuron based on its weighted-adjusted input (Coakley & Brown, 2000). Generally speaking, MLP uses a linear combination function and sigmoid transfer function in the hidden layer (Sarma, 2013). Sigmoid transfer functions include Arc Tangent, Elliot, Hyperbolic Tangent and Logistic, which are S-shaped and have an output values range of -1 to 1. Coakley and Brown (2000) also noticed that none of the research papers they reviewed reported the use of the sine transfer function. Some researchers recommended the logistic transfer function for classification problems involving learning about average behaviour, and the hyperbolic tangent transfer function for prediction issues involving learning about deviation from the average (Coakley & Brown, 2000). Coakley, McFarlane, and Perley (1992) assessed the forecasting performance of three transfer functions (logistic, half-logistic and hyperbolic tangent) when used in a MLP with backpropagation learning algorithm and found that the hyperbolic tangent transfer function offered faster convergence and slightly better predictive accuracy than others. Brown, An, Harris, and Wang (1993) also suggested that it was better to use the hyperbolic tangent transfer function in the hidden layer.

Although there are no clear criteria to select a specific learning rate, momentum rate, transfer function, and error function, considering their relationship can be helpful. The current study adopted a gradient descent backpropagation learning algorithm to deal with the prediction problem. Thus, SSE error function was appropriate and the hyperbolic tangent transfer function was applied to the hidden layer. Coakley et al. (1992) suggested that lower values of learning rate and momentum rate should be set if the hyperbolic tangent transfer function is used. As a starting point, a hybrid optimisation algorithm, including a genetic algorithm, was applied, to look for the optimal set of hyperparameters, which is supposed to be better than grid search or random search.

4.4.3.3 Training process

A data transformation is generally required in order to properly facilitate training of the neural network. The input data for the neural network need to be scaled to match the range of the transfer function through normalisation (midrange) or standardisation (standard deviation). Shanker, Hu, and Hung (1996) evaluated both linear transformation (a range of 0 to 1) and statistical standardisation and found that standardisation improved classification rate with more computation time. For the hyperbolic tangent transfer function applied to the hidden layer, input data will be scaled within the range of -1 to 1. Amongst training samples, transformed input data values near the extreme will be clipped at the upper or lower bound of the transfer function. The clipping procedure in the optimisation task could be used to ensure that the neural network error function does not penalise values outside of the range (Goodfellow et al., 2016). Eventually, any differences between the input data range and range of transfer function would be compensated by the derived weights (Coakley & Brown, 2000). It is recommended that the output data for the neural network be scaled as well, unless the identity transfer function is applied to the output layer. Scaled output data values near the boundary will cause saturation of the network (Coakley & Brown, 2000).

The connection weights are commonly initialised by randomly assigning a smaller number prior to training. Additionally, heuristic approaches and algorithms have been developed to determine the appropriate initial weights and mitigate the risk of a big error and local minima associated with random approaches. Wessels and Barnard (1992) suggested that the initial weights should be within the range of +-3A/sqrt(N), where A is the standard deviation of the inputs to the node and N is the number of weights pointing at the node. Algorithms used to initialise weights are Adaptive and Xavier weight initialisations. Bias on both the hidden layers and output layers is initially introduced to generalise the neural network. Thereafter, the weights are adjusted to minimise the training error each epoch. The neurons with the smallest sum of weights can be removed during training to improve the performance of the neural network due to their least effect on the solution.

The training process will be terminated when a specific threshold is met: (1) convergence occurs; (2) maximum of iterations is reached; or (3) the error on the validation dataset starts to increase. Maximum of iterations is predefined to avoid overtraining where the neural network attempts to exactly fit the set of data points but loses the ability to learn the relationship between those data points (Hecht-Nielsen, 1989). However, if a holdout validation method is used to evaluate the performance of neural networks after each epoch, the training will stop when performance on the validation dataset starts to deteriorate. The network will show signs of overtraining to the training dataset when the error for the validation dataset starts to increase even though the training error rate continues to reduce.

In the current study, input data for the neural network were scaled into the desired range constrained on the hyperbolic tangent transfer function via statistical standardisation. Weights and bias were randomly initialised. Training process was terminated as performance on the validation dataset started to deteriorate. To provide a basis for model comparison with regressions, AIC was calculated for the neural network, as shown in Table 6. A comparison with the AIC values for the regression in Table 5 indicates that the neural network model underperformed in developing a precise expectation for the financial account balances in this study’s implementation.

4.4.3.4 The statistical decision rules

A statistical decision rule was necessary to assess the performance of the neural network method as an analytical procedure. Since output data for the neural network is unbounded, given the use of identity transfer function, the predicted output values derived by the neural network can be directly interpreted. As with regression procedures, the deviation between a predicted and an observed output value can be estimated using Equation 6. After a prediction interval was constructed, the same statistical decision rules which were used to evaluate regression models were applied to the neural network model to be assessed as an analytical procedure.

Table 6 AIC of neural network analysis

Financial account	AIC
Receivables	1173
Inventory	1173
Payables	1173
Sales	1173
Cost of sales	1173
Current assets	1173
Current liabilities	1173

4.5 Experimental procedures

SAS Enterprise Miner 14.3 and SAS Visual Data Mining and Machine Learning were the software programmes used to predict interval dependent variables. Data preparation involved assessing dependent variables’ distributions, identifying outliers, and removing spurious correlated variables. Data from six of the case firms were randomly partitioned to train and validate models with the rule of 80% and 20%. A training set was used to search for parameters and develop models. A validation set was used to evaluate the performance of models and assess the generalisation. Data from the last case firm were used to test the models. Modelling issues, such as autocorrelation, heteroscedasticity and multicollinearity, were considered. The dataset was relatively small, as is the case in most of the studies in the accounting literature, which may present bias and could be considered as a limitation.

The effectiveness of analytical procedure techniques was evaluated in terms of error detection ability. If the analytical procedure technique was efficient, it would not signal the existence of errors in the case where no errors or immaterial errors were seeded. If the analytical procedure technique was reliable, it would signal the presence of seeded material errors for further investigation. An error may occur at any time in a given period in an audit situation. Knechel (1986) and Kinney and Salamon (1982) have demonstrated that analytical procedures are less likely to detect smaller errors spreading over a number of periods. Therefore, an annual material error was seeded into a firm-specific, semi-annual account balance in this study. An error was seeded in either the first half of the year or the second half of the year and did not spread to the other halves of the years. In the same manner as Coakley and Brown (1993), three factors were varied during the simulation. They were the size of the monetary error, the statistical level of confidence placed on the analytical procedures and the sources of material errors. The research approach produced 480 different comparisons (six firms x two halves of year x two inventory conditions x two sources of error x five material amounts x two alpha risk levels).

The size of the monetary error was varied from 0, 0.5, 1.0, 1.5 and 2.0 times the materiality threshold (M). The number of Type I errors was expected to increase when the size of the error grew from 0 to 0.5 times M. The efficient analytical procedure should not signal further investigation because the error was less than M. Generally, the larger errors should be easier to find. Therefore, the number of Type II errors was expected to decrease when the size of the error increased from 1.0 to 2.0 times M.

The width of prediction interval was determined by the statistical level of confidence placed on the analytical procedure (Coakley, 1982; Kinney, 1987; Loebbecke & Steinbart, 1987). The area between upper bound limit and lower bound limit becomes wider if alpha level (α) is low, which should lead to fewer investigation signals. On the other hand, a higher value of α gives a narrow area, resulting in more Type I errors but fewer Type II errors. In this study, α levels of 0.10 and 0.33 were applied.

The two types of alternative seeded error were unrecorded purchase and fictitious sales, respectively. Also, two kinds of inventory condition were considered: no physical count and adjustment to the inventory record after a physical count. In order to fairly compare the financial ratios with the results of Coakley and Brown (1993), the assumption was used that inventory account was adjusted to match the physical count after a physical count of inventory was taken. For the rest of the analysis, the assumption that the inventory accounts were not adjusted according to a physical count was used.

Finally, developed models were also tested with data obtained from the last case firm. Predicted values were compared to recorded values and the discrepancy was assessed based on the investigation rule. Investigation signals were evaluated against the due diligence report produced by a financial analyst in the U.S.

5 RESULTS

The discussion of results is divided into three subsections as follows: (1) financial ratios; (2) comparison of methods; (3) additional analyses through experimenting with the effect of error sizes, statistical confidence levels and error sources, and applying the developed models to score.

5.1 Financial ratios

Results of financial ratios in the current study were compared to the results of Coakley and Brown (1993) and are shown in Table 7. The means and standard deviations of financial ratios were different from the prior study because of the use of the cross-sectional time series data in this study. Type I error rates when no error was seeded into accounts and Type II error rates when an error twice materiality was seeded into accounts were compared. The prediction period was the financial year 2016.

Table 7 Comparison of financial ratio analysis

	Coakley & Brown (1993)					Case study				
Financial ratios	Mean	Std. Dev.	Type I	Type II	Total	Mean	Std. Dev.	Type I	Type II	Total
Receivables turnover	0.35	0.03	0.26	0.72	0.98	11.21	15.20	0.00	1.00	1.00
Inventory turnover	0.84	0.20	0.59	0.40	0.99	2.16	1.82	0.21	0.79	1.00
Cost of sales ratio	0.71	0.02	0.56	0.00	0.56	0.61	0.09	0.13	0.87	1.00
Accruals ratio	1.55	0.10	0.44	0.42	0.86	0.67	0.40	0.29	0.67	0.96
Quick ratio	0.72	0.04	0.72	0.26	0.98	0.82	0.70	0.04	0.96	1.00

Coakley and Brown (1993) showed different results compared with those of Kinney (1987), even though the financial ratios had similar means and standard deviations across the two studies. They concluded that “it was much more difficult to distinguish the fluctuations caused by seeded errors from those that normally occurred in the financial data of the firm” (p. 13). Therefore, the current study was expected to have very different results from Coakley and Brown (1993) because multiple firms were involved. When compared with the results of Coakley and Brown (1993), applying those financial ratios to the data in the current study produced higher Type II error rates and lower Type I error rates. The quick ratio is presented solely for comparison purposes because Coakley and Brown (1993) did not discuss it in detail.

For both Coakley and Brown (1993) and the current study, applying Equation 2 from Chapter 3 resulted in a relatively low materiality threshold, which was approximately 0.01% of annual revenue of all firms for the prediction period. The resulting value for a material error was about a half of the value derived from the formula recommended in general materiality guidelines (1% of annual revenue[4]) if the smallest firm in this study is considered. Loebbecke and Steinbart's (1987) analysis showed that the average prediction errors generally exceeded the average materiality measures (54 out of 70 cases) and provided evidence that the signalling abilities of an analytical procedure may be tied to the relationship of prediction errors to the level of materiality. As a result, the lower materiality threshold used would explain relatively higher Type II error rates as it should be easier to detect larger errors. As Coakley and Brown (1993) correctly pointed out, materiality does not affect the decision rule causing the Type I error for the financial ratios. The difference in the Type I error rates could be explained by the idea that the efficiency of financial ratios as an analytical procedure would be improved, given the cross-sectional comparison.

In the current study, the decision rule concerned the comparison between the value of the financial ratio of a firm at a particular time with the average of the financial ratio values of all firms from the prior audit period. The Type I errors were due to the natural variability in the account balances. Figure 3 shows the averaged ratios for all case firms over the period. In Appendix II, the comparison between individual firm's ratios and averaged ratios is provided. High variability in the calculated inventory turnover ratio values would yield false investigation signals. However, the values of the cost of sales ratio appeared to be consistent over the period. The very low variance would lead to small fluctuations to trigger investigation signals even if there were no material errors in the account. The consistently low

[4] Institute of Chartered Accountants in England and Wales (ICAEW) https://www.icaew.com/-/media/corporate/files/technical/iaa/materiality-in-the-audit-of-financial-statements.ashx

variation potentially equips the cost of sales ratio with a better capability of detecting seeded errors, which would result in very low Type II error rates compared to other financial ratios. Offsetting this is the fact that this ratio consists of a very large numerator and very large denominator (e.g., cost of sales and sales) so material errors could be swamped by the size of each of the ratios' variables (Kinney, 1987). Although the values of the accruals ratio appeared consistent over the period, the wide variation across the individual values resulted in many false investigation signals.

Loebbecke and Steinbart (1987) and Coakley and Brown (1993) found that financial ratios do not reliably indicate the absence of errors. However, the comparison showed that the Type I error rates significantly dropped from the Coakley and Brown (1993) study, which used a single firm, to the current study involving multiple firms. The result suggests that the financial ratios could be much more effective analytical procedures if industry peers' data are used and the materiality threshold is predetermined appropriately. The result is also consistent with the suggestion of prior research and auditing standards which encourages auditors to compare the firm with its industry peers when applying financial ratios as an analytical procedure.

5.2 Comparison of methods

The first two hypotheses, H_1 and H_2, posit:

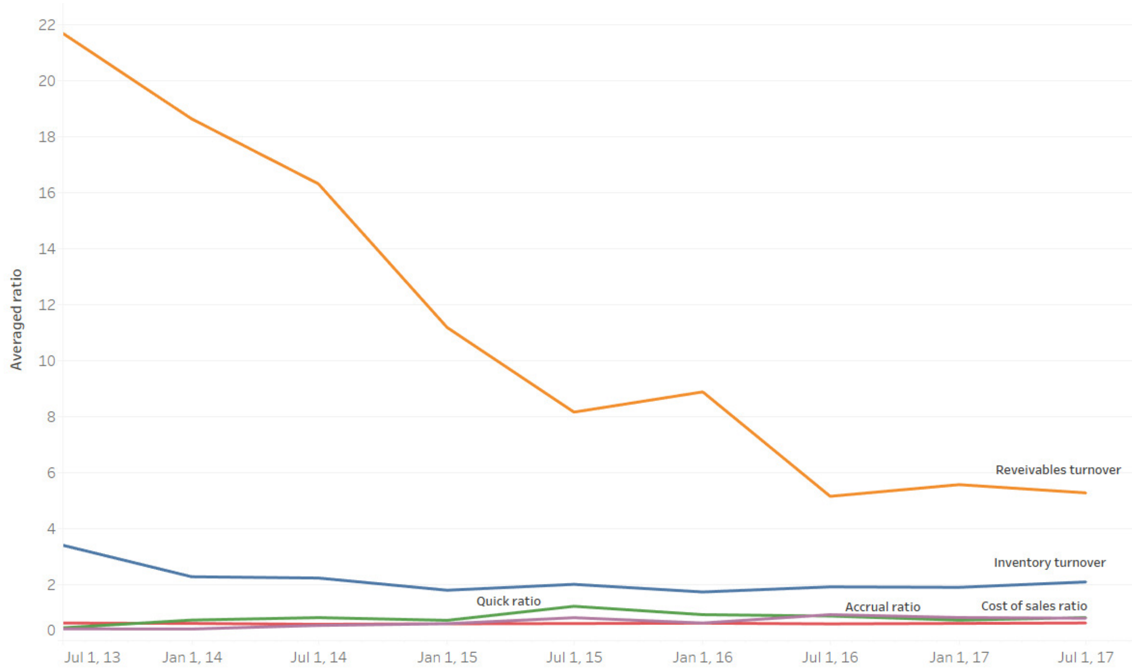
H_1: The neural network approach is more effective than the two alternative approaches.

H_2: None of the three approaches provide an improvement in overall effectiveness over the benchmark of 1.0.

A comparison of the average results across the three factors are presented in Table 8 and depicted in Figure 4 for the financial ratio, regression and neural network methods, based on both inventory assumptions. In Figure 4, NA denotes the assumption that no physical count was taken, ADJ is for the assumption that inventory values were adjusted to the physical count, and the solid line in the upper right-hand corner of the graph represents the 1.0 benchmark (Loebbecke & Steinbart, 1987). An analytical procedure method will be effective if the sum of Type I and Type II error rates is less than 1.0.

Compared to the prior study, financial ratio procedures in the current study resulted in a much lower Type I error rate. Reasons for that were explained in subsection 5.1. Furthermore, when larger errors were seeded, the financial ratio procedure would become slightly more effective based on the average rates that are the total of Type I and Type II error rates, as its Type II error rate quickly decreased. In terms of the sensitivity of ratios to the two inventory assumptions, Coakley and Brown (1993) noticed that a difference occurred in the effectiveness of the financial ratios because the assumption of adjusting the inventory to the physical count helped the inventory turnover ratio lower the Type I error rate, and the assumption of not adjusting inventory to physical count removed the effect on the cost of sales ratio, to increase the Type II error rate. However, individual firms had a varying stock management efficiency and gross profitability in the current study, which minimised the effect of varying the inventory assumption. Consequently, the current study produced similar Type I and Type II error rates for both inventory assumptions.

Figure 3 Financial ratios calculated from audited balances



For the regression and neural network methods, results in the current study did not show an insignificant increase in overall effectiveness relative to ratio analysis that was mentioned by Coakley and Brown (1993). The neural network method in the current study produced a lower Type II error rate (enhanced reliability) and a higher Type I error rate (reduced efficiency), as compared to the other approaches, a result that is similar to that of Coakley and Brown (1993). However, based on the sum of Type I and Type II error rates, Coakley and Brown (1993) concluded that the neural network approach was slightly more effective than other methods, which was not the case in the current study. The total error rate for each analytical procedure method approximately equalled the benchmark of 1.0. Therefore, none of the three approaches provided an improvement in effectiveness over a purely random process.

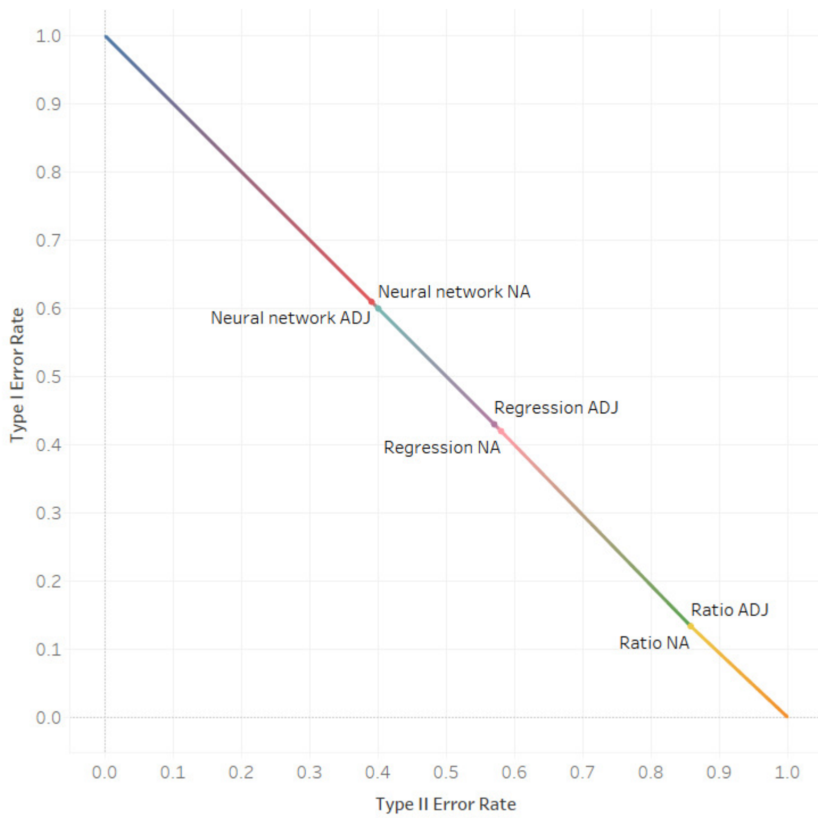
H_2 is provisionally supported, but not H_1. The results were not far different from the conclusion of Coakley and Brown (1993). To examine possible factors which were attributable to the results, the next subsection will analyse the performance of the three methods in three aspects: the effect of error size, the effect of statistical level of confidence, and the effect of source error.

Table 8 Comparison of analytical procedure methods

	Coakley & Brown (1993)			Case study		
Analytical procedure methods	Type I	Type II	Total	Type I	Type II	Total
Financial ratio – NA	0.45	0.43	0.88	0.13	0.86	0.99
Regression – NA	0.35	0.49	0.84	0.42	0.58	1.00
Neural network - NA	0.57	0.25	0.81	0.60	0.40	1.00
Financial ratio – ADJ	0.40	0.55	0.95	0.13	0.86	0.99
Regression – ADJ	0.40	0.46	0.86	0.43	0.57	1.00
Neural network – ADJ	0.52	0.30	0.82	0.61	0.39	1.00

Note. NA: No physical count was taken*;* ADJ: Inventory values were adjusted to the physical count.

Figure 4 Comparison of analytical procedure methods



5.3 Additional analyses

Error rates for the three methods were influenced by three varying factors in the experiment. The performance of the analytical procedures was further evaluated by applying the developed models to the testing dataset. Accordingly, five hypotheses were examined and subsequently rejected if a difference was observed. Results are shown in the following four subsections: (1) effect of error size; (2) effect of statistical level of confidence; (3) effect of source error; and (4) practical application to holdout case firm.

5.3.1 Effect of error size

The plausibility of H_3 and H_4 were studied:

H_3: The three approaches have increased error rates when the size of the immaterial error increases.

H_4: The three approaches have reduced error rates when the size of the material error increases.

The impact of the size of the error seeded into the financial accounts on the Type I and Type II error rates is shown in Table 9 and Figure 5. The values depicted were computed by averaging over the two sources of error and two statistical levels of confidence. In Coakley and Brown's (1993) work, when seeding an immaterial error and increasing its size, the regression and neural network methods resulted in higher Type I error rates, but the financial ratio approach led to a lower Type I error rate, most likely due to the fluctuating nature of the data over the test period. In contrast, the current study showed that the three methods had almost unchanged Type I error rates. Therefore, H_3 was not supported.

When larger seeded errors were considered, all procedures became more reliable as larger errors were easier to detect. If the size of the error was raised from 1 time to 30 times M, Type II error rates for the financial ratio procedure fell from 0.86 to 0.78 (NA) and 0.68 (ADJ). As shown in Appendix III, if 1, 1.5 and 2 times M were replaced by 15, 20 and 30, its Type II error rate, for the effect of alpha risk (0.33), and sources of error (purchases not recorded, P), respectively, dropped to 0.76 (NA) and 0.68 (ADJ) and 0.80 (NA) and 0.70 (ADJ). However, surprisingly, Type II error rates for the regression and neural network methods slightly changed. H_4 was supported, although the reduction of error rates for both the regression and neural network approaches was not very obvious in this study's implementation.

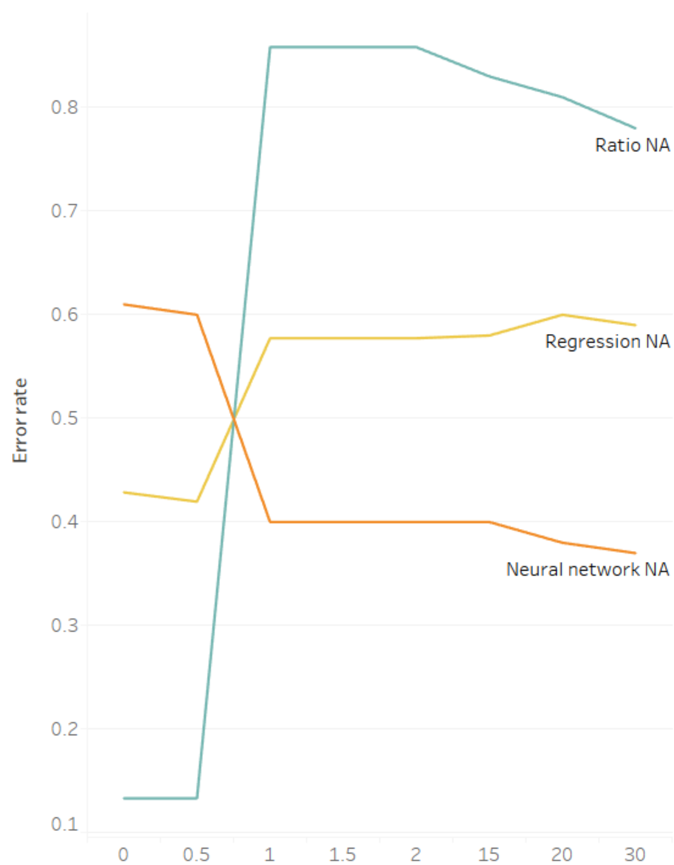
The sum of Type I and Type II error rates for the three methods approximately equalled the benchmark of 1.0. However, if the size of the error increased up to 30 times M, the total error rate almost did not change for the regression and neural network approaches, while it did for the financial ratio approach. In contrast to the results of Coakley and Brown (1993), there seemed to be a difference in the improvement across the three methods.

Table 9 Effect of error size on error rates

	Financial ratio		Regression		Neural network	
Error sizes	NA	ADJ	NA	ADJ	NA	ADJ
0 times M	0.13	0.13	0.43	0.43	0.61	0.61
0.5 times M	0.13	0.13	0.42	0.43	0.60	0.61
1 times M	0.86	0.86	0.58	0.57	0.40	0.39
1.5 times M	0.86	0.86	0.58	0.57	0.40	0.39
2 times M	0.86	0.86	0.58	0.57	0.40	0.39
15 times M	0.83	0.78	0.58	0.58	0.40	0.41
20 times M	0.81	0.75	0.60	0.58	0.38	0.40
30 times M	0.78	0.68	0.59	0.58	0.37	0.40

Note. NA: No physical count was taken; ADJ: Inventory values were adjusted to the physical count.

Figure 5 Effect of error size on Type I and Type II error rates



5.3.2 Effect of statistical level of confidence

This subsection examines H_5 which assessed the effect of statistical level of confidence on error rates:

> H_5: The neural network approach has less trade-off between Type I and Type II errors than alternative approaches when varying the alpha risk.

The effect of the different statistical level of confidence (alpha risk level) is shown in Table 10 and Figure 6. The calculated values were averaged over the seeded error size and the source of error. The prediction interval was derived from the alpha risk and the variation in the predicted values. The smaller variations produced tighter prediction intervals. Compared to regressions, the neural network method produced higher AIC and variation in this implementation. As a result, the preliminary conclusion of Coakley and Brown (1993), that the neural network approach was less sensitive to varying the alpha risk, did not appear in the current study. Instead, the results in Figure 6 show a similar sensitivity to varying the alpha risk for the three methods.

The results were consistent with Coakley and Brown (1993) and showed that all approaches had approximately the same overall effectiveness when the alpha risk was increased. The slope of all the lines roughly overlapped with the 1.0 benchmark line, which indicates that there was a trade-off between Type I and Type II errors as the alpha risk was increased from 0.10 to 0.33. As with the above-mentioned sensitivity, the results in Figure 6 show a similar trade-off available for the three methods when varying the alpha risk. As a result, H_5 was not supported.

Table 10 Effect of statistical level of confidence on error rates

	Financial ratio			Regression			Neural network		
Confidence levels	Type I	Type II	Total	Type I	Type II	Total	Type I	Type II	Total
0.10 – NA	0.08	0.90	0.98	0.36	0.64	1.00	0.55	0.46	1.01
0.33 – NA	0.18	0.82	1.00	0.48	0.52	1.00	0.65	0.35	1.00
0.10 – ADJ	0.08	0.90	0.98	0.36	0.64	1.00	0.56	0.44	1.00
0.33 – ADJ	0.18	0.82	1.00	0.49	0.51	1.00	0.65	0.35	1.00

Note. NA: No physical count was taken; ADJ: Inventory values were adjusted to the physical count.

Figure 6 Effect of alpha risk level on Type I and Type II error rates

5.3.3 Effect of source of error

This subsection examines H_6, which looked at the effect of source of error on error rate:

> H_6: The regression and neural network approaches react similarly to the purchases not recorded (P) error, but not to the fictitious sales (S) error.

The effect of the source of material error on the Type I and Type II error rates is presented in Table 11 and Figure 7. The computed values were averaged across the size of error seeded and the alpha risk levels. In Figure 7, the range of error rate for each method becomes a point below or on the benchmark line 1.0, which implies that the three methods had similar overall effectiveness across both error types, but with a different trade-off between the Type I and Type II error rates.

The regression and neural network methods similarly reacted to both the purchases not recorded (P) error and the fictitious sales (S) error, which was different from Coakley and Brown (1993) stating that the two methods seemed to have similar overall effectiveness with P, but a completely opposite effect with S. Consequently, H_6 was not supported. Table 12 and Figure 8 show that the ability of both methods to signal errors in these financial accounts were the same with a different trade-off between the Type I and Type II error rates because they produced a similar overall effectiveness in signalling investigations.

For the fictitious sales error, the sales, cost of sales, inventory and receivables accounts would be directly affected. The average Type II error rate for the regression method is driven by the high Type II error rates in the sales and inventory accounts, offset by the lower Type II error rates in the cost of sales and receivables accounts. On the other hand, the average Type I error rate for the neural network method is due to the higher Type I error rates in the sales, cost of sales and receivable accounts, adjusted by the lower Type I error rate in the inventory account.

For the purchases not recorded error, the cost of sales, inventory and payables accounts would be strongly affected. The average Type II error rate for the regression method is driven by the high Type II error rates for the inventory and payables accounts, which is offset by the high Type I error rates for the cost of sales account. The average Type I error rate for the neural network method is due to the higher Type I error rates in the cost of sales and payables accounts, adjusted by the lower Type I error rate in the inventory account.

Table 11 Effect of source of error on error rates

	Financial ratio			Regression			Neural network		
Sources of error	Type I	Type II	Total	Type I	Type II	Total	Type I	Type II	Total
Purchases – NA	0.13	0.86	0.99	0.42	0.58	1.00	0.61	0.40	1.01
Sales – NA	0.13	0.86	0.99	0.42	0.58	1.00	0.60	0.41	1.01
Purchases – ADJ	0.13	0.86	0.99	0.42	0.58	1.00	0.61	0.40	1.01
Sales – ADJ	0.13	0.86	0.99	0.43	0.57	1.00	0.61	0.39	1.00

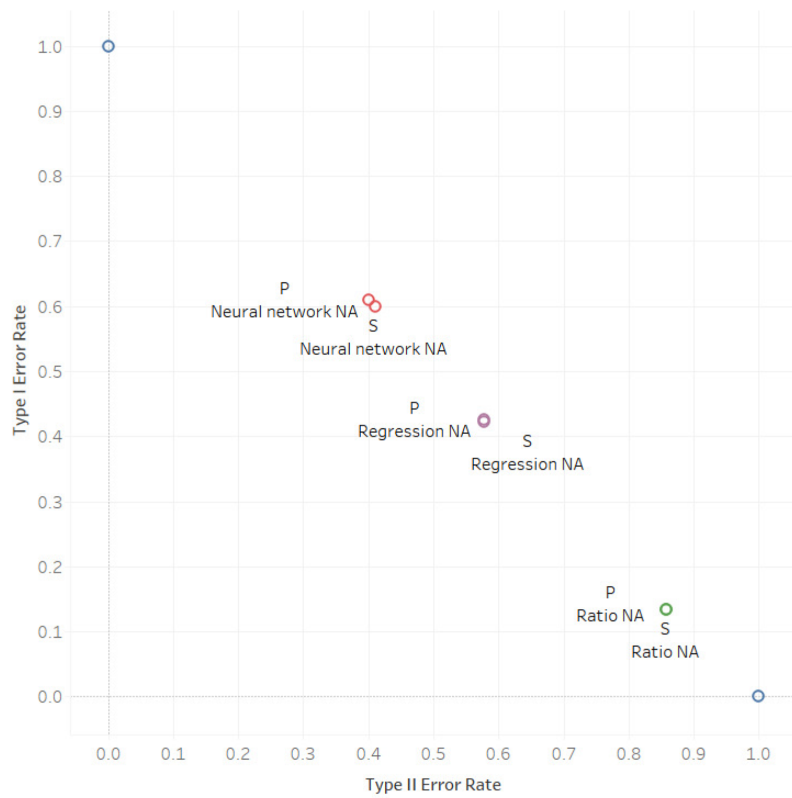
Note. NA: No physical count was taken; ADJ: Inventory values were adjusted to the physical count.

Table 12 Error rates by individual financial account

Individual account	Regression			Neural network		
	Type I	Type II	Total	Type I	Type II	Total
Inventory – NA	0.31	0.71	1.02	0.36	0.65	1.01
Cost of sales – NA	0.87	0.13	1.00	0.79	0.21	1.00
Sales – NA	0.41	0.60	1.01	0.74	0.27	1.01
Receivables (AR) – NA	0.67	0.31	0.98	0.74	0.26	1.00
Payables (AP) – NA	0.46	0.54	1.00	0.65	0.37	1.02
Current assets (CA) – NA	0.17	0.83	1.00	0.21	0.79	1.00
Current liabilities (CL) – NA	0.08	0.92	1.00	0.71	0.29	1.00
Inventory – ADJ	0.33	0.67	1.00	0.37	0.63	1.00
Cost of sales – ADJ	0.87	0.15	1.02	0.79	0.21	1.00
Sales – ADJ	0.41	0.60	1.01	0.75	0.25	1.00
Receivables (AR) – ADJ	0.67	0.31	0.98	0.75	0.24	0.99
Payables (AP) – ADJ	0.46	0.54	1.00	0.66	0.35	1.01
Current assets (CA) – ADJ	0.17	0.83	1.00	0.21	0.79	1.00
Current liabilities (CL) – ADJ	0.08	0.92	1.00	0.71	0.29	1.00

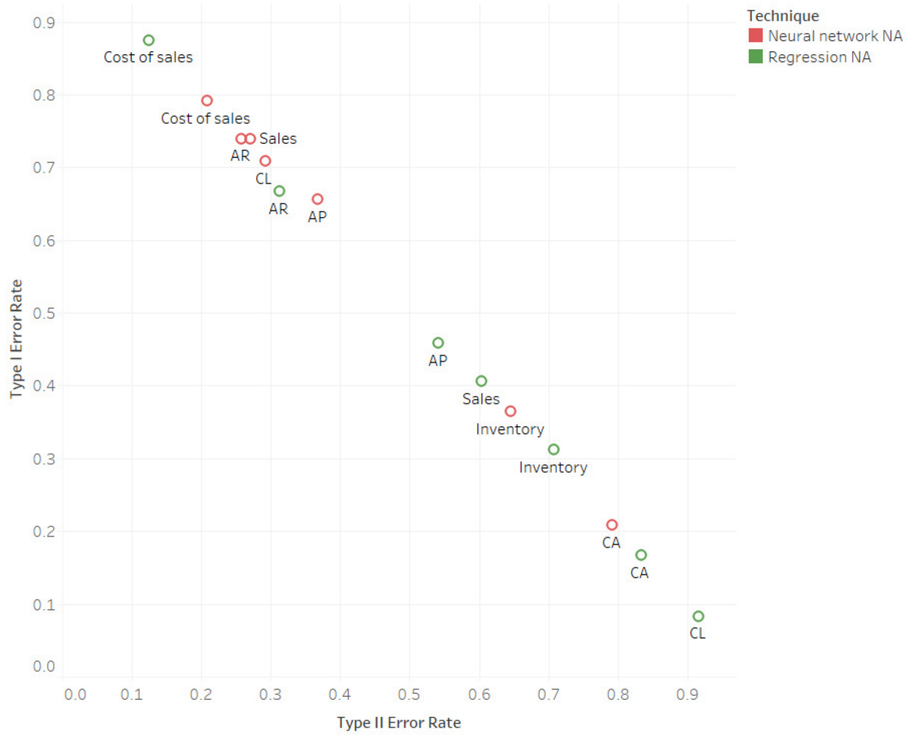
Note. NA: No physical count was taken; ADJ: Inventory values were adjusted to the physical count.

Figure 7 Effect of source of error on Type I and Type II error rates



Note. P: The purchases not recorded error; S: The fictitious sales error.

Figure 8 Regression and neural network methods error rates by individual financial account



5.3.4 Practical application of methods to holdout case

The developed models were applied to the testing dataset derived from the holdout case firm (a listed Chinese dairy company) to develop expectations for the financial year 2016. The difference between the expected and recorded value was assessed based on the statistical rule and then investigation signals were generated for five ratios and seven financial accounts. Further, results of the three approaches were compared with the due diligence report of the financial analyst.

Before showing the diagnostic results of each of these methods, a brief background of the firm is needed. In 2016, a financial analyst based in the U.S. shorted the firm and alleged that its financial report was fraudulent and value was close to zero. Red flags identified by the analyst were the overstated growth rate, fictitious sales, understated cost of sales, overstated gross profit margin, overstated property, plant and equipment, non-existing cash, suspicious related party transactions, overstated fair value of biological assets, unreasonable prepayment, the risk of net realizable value (NRV) of inventory and high liquidity risk. In early 2017, through several company announcements, the firm admitted the discrepancy of cash RMB 2.4 billion, defaulted debts and its missing head of treasury, and engaged one of the Big Four audit firms to provide forensic services. All information suggested that there was a reasonable chance that sales and cost of sales accounts could be subject to material misstatements, due to either error or fraud. Hence, investigation signals were expected for financial accounts, such as sales, cost of sales, inventory, receivables, payables, current assets, and current liabilities. As a result, the following hypothesis was derived:

H_7: The neural network approach generates investigation signals for all seven financial accounts to support assertions of a financial analyst.

Results are shown in Table 13. The neural network met expectations and agreed with the due diligence report findings of the financial analyst, probably because this approach produced the lower Type II error rate as compared to the financial ratio and regression approaches. Specifically, for financial ratio procedures, investigation signals could be obtained for sales, cost of sales, inventory, receivables and payables accounts from inventory turnover, cost of sales ratio and quick ratio. For the regression method, investigation signals were derived from sales, cost of sales, inventory and receivables accounts. Noticeably, when the neural network approach was applied, all seven accounts signalled the need for further investigation. Hence, H_7 was fully supported.

Table 13 The developed models applied for testing

Financial ratios and individual account	Financial ratio		Regression		Neural network	
	Difference*	Signal	Difference	Signal	Difference	Signal
Receivables turnover	0.04	-	-	-	-	-
Inventory turnover	1.73	Yes	-	-	-	-
Cost of sales ratio	-7.55	Yes	-	-	-	-
Accruals ratio	-0.93	-	-	-	-	-
Quick ratio	6.55	Yes	-	-	-	-
Inventory	-	Yes	7.41	Yes	9.02	Yes
Cost of sales	-	Yes	-5.88	Yes	-17.90	Yes
Sales	-	Yes	-21.15	Yes	-20.53	Yes
Receivables	-	Yes	-18.44	Yes	-4.55	Yes
Payables	-	Yes	-0.64	-	5.03	Yes
Current assets	-	-	-0.59	-	3.29	Yes
Current liabilities	-	-	-0.08	-	4.30	Yes

* Difference: The standardised discrepancy between the expected and recorded amount.

6 DISCUSSION AND LIMITATIONS

After developing precise expectations, the overall research hypothesis was that the neural network analytical procedure would be better at detecting errors than would the financial ratio and regression approaches or a purely random procedure. The results suggest that the neural network method's predictive performance was not superior to regression methods based on the AIC model comparison in this study's implementation. However, this finding was not conclusive, as many other factors and techniques, which may significantly improve the performance of neural networks, were not covered in the current study. The results also suggest that the neural network method was not slightly more effective than alternative methods in terms of overall error detection performance (combined error rate). Consistent with Coakley and Brown's (1993) findings for a single firm, none of the three methods provided a significant improvement in effectiveness over a purely random procedure, as the sum of Type I and Type II error rates for each method approximately equalled the benchmark of 1.0.

In order to reach these conclusions, the performance of the three methods was evaluated by varying error sizes, statistical confidence levels and error sources. Compared to the experimental results of Coakley and Brown (1993), the current study had different findings. When the size of the error increased, the Type I error rate for all approaches remained nearly unchanged. The Type II error rate for the financial ratio method reduced significantly, while the Type II error rate for the regression and neural network methods slightly changed. Thus, there appeared to be different improvements in effectiveness across the three methods when changing the size of error, since the financial ratio method became more effective than alternative methods if larger errors were seeded. When varying the alpha risk, all approaches had similar trade-offs between the Type I and Type II error rates and overall effectiveness. When investigating sources of error, the regression and neural network approaches reacted similarly to both errors and there appeared to be similar overall effectiveness across all approaches. Finally, the neural network approach met expectations and supported the assertions of the financial analyst when the developed models were applied to the testing dataset.

It is noteworthy that both Coakley and Brown (1993) and the current study report that the neural network approach did enhance reliability by producing a lower Type II error rate but reduced efficiency by yielding a higher Type I error rate. Also, Type II error rates in both studies were below 0.50, which is considered to be a control point to determine if analytical procedures are effective. As a result, the findings imply that neural networks may be effective in detecting material errors when they indeed exist but could also expose auditors to unnecessary investigations. Auditors must select an expectation model in light of their relative risk preference. Normally, they give Type II errors more weight than Type I errors, due to high litigation cost involved, and attempt to reduce detection risk in accordance with audit risk model. If an auditor's objective is not overall performance, but instead to minimise the Type II error rate (which is concluding that an account is not in material error when it is), the results suggest that the neural network approach is useful in practice.

From another perspective, the following two findings are noted. The current study extends the work of Coakley and Brown (1993) by incorporating both multiple firm year data and exogenous variables. Prior research suggested that the use of external data (peer, industrial and economic data) can improve the precision of expectation models and the effectiveness of analytical procedures (Chen & Leitch, 1998; Hoitash et al., 2006; Lev, 1980; Loebbecke & Steinbart, 1987; Neter, 1980; Wild, 1987). When compared to Coakley and Brown (1993), the financial ratio procedure in the current study resulted in a significantly lower Type I error rate. This observation is consistent with prior research findings and the suggestion of auditing standards. In addition, the results clearly support the conclusion of Loebbecke and Steinbart (1987) and Pany and Wheeler (1992) that, while these methods may be good at spotting fluctuations resulting from the presence of material monetary errors in the account balances, they do not reliably indicate the absence of material monetary errors.

The current study had several limitations. First, it used a relatively small dataset, as is the case in most of the studies in the accounting literature, and this will have reduced the statistical power of tests and may limit the generalisability of the study's findings. Generally, techniques such as the k-fold cross-validation and jackknife method (Coakley, 1995) are used to compensate for this issue. However, the current study did not apply any of these because the statistical software used did not provide obvious options in the graphical user interface (GUI). Second, the formula (Equation 2 from Chapter 3) used to determine the materiality threshold resulted in a very low value which may lead to exaggerated Type I errors but lower Type II errors. Normally, auditors have access to information about the characteristics and the background of their clients. With this valuable information, they should be able to establish more appropriate materiality thresholds that are better than those determination formulas used in this research.

Third, the current study seeded errors in one time period across multiple companies and did not assess the performance of analytical procedures when small errors were dispersed throughout the year making them difficult to detect. Under such circumstances, it would be hard for any procedure in the research to distinguish between small errors and normal fluctuations in account balances. This represents a challenge in determining an appropriate procedure, as small errors can accumulate into a material error. Fourth, the neural network approach produced less precise expectations than the regression approach in this study's implementation, which in turn could affect its Type II error rate. The results showed that the neural network method seemed preferable to the regression method in term of Type II error rate. However, the possibility cannot be ruled out that this was actually due to the poor prediction performance of the neural network model, resulting in prediction errors so large that a material error was so often concluded and there was little chance that a material error would be missed. Arguably, the successful application of the neural network to the holdout case firm could ease that concern.

7 CONCLUSIONS

Neural networks have been proven to possess superior predictive performance with less overall variation in the predicted values from the recorded amount and pattern recognition ability to analyse complex relationships. Previous research suggests that there is a need to develop sophisticated expectation models like neural networks that can improve the efficiency and effectiveness of analytical procedures. Coakley and Brown (1993) assessed the performance of a neural network as an analytical procedure for a single case firm and felt that their results were inconclusive. The current study extended the work of Coakley and Brown (1993) by using multiple case firms and exogenous variables. To evaluate the effectiveness of neural networks as an analytical procedure, it was necessary to test whether the predictive ability and overall error detection performance of neural networks was better than the financial ratio and regression approaches.

This study observed that the neural network approach was not slightly more effective than the alternative approaches and none of the three methods provided an improvement in effectiveness over a purely random process because the sum of Type I and Type II error rates were approximately the same and equal to the benchmark 1.0. These results were not far different from the conclusion of Coakley and Brown (1993). Furthermore, this research provided evidence that the neural network approach had better reliability (lower Type II error rates) but lower efficiency (higher Type I error rates) compared to the financial ratio and regression approaches, which concurs with Coakley and Brown’s (1993) findings.

From the auditor’s perspective, the costs associated with the ease of use of these expectation models against the potential costs due to the decision risk must be balanced. In general, auditors may select the appropriate approach which will result in desired risks being close to planned levels. The results here showed that the financial ratio method can enhance audit efficiency through reducing excessive audit effort when no material error is present, while still providing the satisfactory assurance of detecting a material error. The financial ratio approach exhibits a needed investigation-signalling ability and is an economical way to perform analytical procedures. The results also suggested that the neural network approach was a more effective way to develop expectation models which are used to alert auditors to material errors and unusual economic events when they are present. Therefore, these findings should help auditors to select an appropriate analytical procedure method to develop an expectation model and detect material misstatements due to error or fraud.

Due to the expensive litigation costs involved, it is normally assumed that auditors would first select the analytical procedures that achieve a specified level of assurance that a material error will be detected, then select a procedure which minimises their extra audit effort. Hence, the results in the current study showed that neural networks were clearly favourable if auditors want to reduce the risk of concluding that a financial account is not in material error when it is in material error. In general, neural networks seem to have the potential of increasing the effectiveness and the efficiency of analytical procedures. As a result, the thesis holds a same opinion as Coakley and Brown (1993) that it is still "worth pursuing the question of whether neural networks are useful as an analytical procedure in auditing" (p. 20).

Neural networks provide a promising alternative analytical procedure even though they may not, in the end, completely replace professional judgment and other methods. The future of neural networks in auditing will be brighter as more research effort is devoted to this area. Future research is needed on the following aspects. First, the quality of the dataset primarily determines the performance of neural networks because of the well-known adage – garbage in and garbage out. The dataset used for training could be expanded to include the entire dairy industry in a particular country since the neural network is hungry for data. The size and the economic stability environment of the companies should be taken into consideration when deciding if a firm is selected for a study. While preparing the dataset in such an information-rich environment, Big Data analytics (Appelbaum et al., 2017; Cao, Chychyla, & Stewart, 2015) can perhaps be used to identify highly relevant peer-based metrics and business patterns and trends to include in the dataset for further analysis. Various data transformation techniques offered by different software packages should be explored.

Second, the application of alternative neural network models to analyse the complicated patterns associated with financial data and precise expectations should be investigated in the auditing domain. For instance, different neural network architectures, learning algorithms, optimisation techniques, activation functions and error functions should be explored. Issa, Sun, and Vasarhelyi (2016) see even further possibilities, beyond the traditional neural networks, in artificial intelligence in auditing and argue that well-trained, deep-learning models enable an auditor to analyse structured or unstructured data without human intervention. Future research perhaps could explore deep learning (convolutional neural networks, CNN) as an analytical procedure capable of extracting the most useful features from the dataset to automatically analyse financial statements. This could also respond to Coakley and Brown's (1993) call to investigate whether the neural network approach could be used to analyse complex patterns in financial accounts in order to identify the source of material error, since no known forecasting method is capable of effectively separating fluctuations caused by errors from the normal fluctuations present in some financial datasets.

Finally, future research ought to compare neural networks with other methods in performing analytical procedures, such as modern statistical models and alternative machine-learning approaches. Given the fact that advanced systems and analytics have various strengths and disadvantages, hybrid models combining neural networks with other models may have better adaptability, predictive performance and error detection ability than individual models separately. Hybrid models and their efficacy should be explored and evaluated.

REFERENCES

Association of Certified Fraud Examiners (ACFE). (2016). *Report to the nations on occupational fraud and abuse*. Retrieved from https://www.acfe.com/rttn2016/docs/2016-report-to-the-nations.pdf

Association of Chartered Certified Accountants (ACCA). (2016). Analytical procedures. Retrieved from http://www.accaglobal.com/an/en/student/exam-support-resources/professional-exams-study-resources/p7/technical-articles/analytical-procedures.html

Akresh, A. D., & Wallace, W. A. (1982). *The application of regression analysis for limited review and audit planning*. Paper presented at the Symposium on Auditing Research IV, Urbana-Champaign, IL.

Alles, M. G. (2015). Drivers of the use and facilitators and obstacles of the evolution of big data by the audit profession. *Accounting Horizons, 29*(2), 439–449.

Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking & Finance, 18*(3), 505–529.

American Institute of Certified Public Accountants (AICPA). (1988). *Analytical procedures. Statement on auditing standards No. 56*. New York, NY: Author.

American Institute of Certified Public Accountants (AICPA). (1995). *General update on economic, accounting and auditing matters*. *Audit risk alert*. New York, NY: Author.

American Institute of Certified Public Accountants (AICPA). (1998). *The auditor's use of analytical procedures, updated August 15, 1999*. *Practice Alert 98-1*. New York, NY: Author.

American Institute of Certified Public Accountants (AICPA). (2004). *Audit and accounting guide: Analytical procedures, update*. New York, NY: Author.

Anandarajan, M., & Anandarajan, A. (1999). A comparison of machine learning techniques with a qualitative response model for auditor's going concern reporting. *Expert Systems with Applications, 16*(4), 385–392.

Appelbaum, D., Kogan, A., & Vasarhelyi, M. (2016). *Analytics in external auditing: A literature review*. Newark, NJ: Rutgers University.

Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2017). Big data and analytics in the modern audit engagement: Research needs. *Auditing: A Journal of Practice & Theory, 36*(4), 1–27.

Arlington, E., Hillison, W, & Icerman, R. (1982). Relative effectiveness and efficiency measures of analytical review procedures. Unpublished manuscript.

Armstrong, W. W., Dwelly, A., Liang, J. D., Lin, D., & Reynolds, S. (1991). Learning and generalization in adaptive logic networks. In T. Kohonen, K. Mäkisara, O. Simula, & J. Kangas (Eds.), *Artificial neural networks* (Vol. 2, pp. 1173–1176). Amsterdam, The Netherlands: Elsevier.

Bai, B., Yen, J., & Yang, X. (2008). False financial statements: Characteristics of China's listed companies and CART detecting approach. *International Journal of Information Technology & Decision Making, 7*(2), 339–359.

Bell, T., Peecher, M., & Solomon, I. (2005). *The 21st century public company audit: Conceptual elements of KPMG's global audit methodology*. Montvale, NJ: KPMG.

Bermúdez, L., Pérez, J., Ayuso, M., Gómez, E., & Vázquez, F. (2008). A Bayesian dichotomous model with asymmetric link for fraud in insurance. *Insurance: Mathematics and Economics, 42*(2), 779–786.

Biggs, S. F., & Wild, J. J. (1984). A note on the practice of analytical review. *Auditing: A Journal of Practice & Theory*, *Spring,* 68–79

Bishop, C. M. (1995). *Neural networks for pattern recognition.* Oxford, UK: Oxford University Press.

Blocher, E., Krull, G. W., Jr., Tashman, L. J., & Yates, S. V. (2002). Updating analytical procedures. *The CPA Journal*, *72*(11), 36–41.

Blocher, E., & Patterson, G. F., Jr. (1996). The use of analytical procedures. *Journal of Accountancy*, *181*(2), 53–57.

Bose, I., & Wang, J. (2007, August). *Data mining for detection of financial statement fraud in Chinese companies*. Paper presented at the International Joint Conference on e-Commerce, e-Administration, e-Society, and e-Education, Hong Kong.

Bridle, J. S. (1990). Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. In F. F. Soulié & J. Hérault (Eds.), *Neurocomputing* (Vol. 68, pp. 227–236). Berlin, Heidelberg: Springer.

Brown, M., An, P., Harris, C., & Wang, H. (1993). How biased is your multi-layered perceptron? (pp. 507–511). Hillsdale, NJ: Lawrence Erlbaum.

Cao, M., Chychyla, R., & Stewart, T. (2015). Big data analytics in financial statement audits. *Accounting Horizons, 29*(2), 423–429.

Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010). Detecting management fraud in public companies. *Management Science, 56*(7), 1146–1160.

Chen, Y., & Leitch, R. A. (1998). The error detection of structural analytical procedures: A simulation study. *Auditing, 17*(2), 36–70.

Cho, S., & Lew, A. Y. (2000). Analytical review applications among large audit firms in Hong Kong. *Managerial Auditing Journal, 15*(8), 431–438.

Coakley, J. R. (1982). *Analytical review: A comparison of procedures and techniques used in auditing* (Unpublished doctoral dissertation). University of Utah, Salt Lake City, Utah.

Coakley, J. R. (1995). Using pattern analysis methods to supplement attention directing analytical procedures. *Expert Systems with Applications, 9*(4), 513–528.

Coakley, J. R., & Brown, C. E. (1991). Neural networks for financial ratio analysis. In J. Liebowitz (Ed.), *Proceedings of The World Congress on Expert Systems* (Vol 1., pp. 132–139). New York, NY: Pergamon Press.

Coakley, J. R., & Brown, C. E. (1993). Artificial neural networks applied to ratio analysis in the analytical review process. *Intelligent Systems in Accounting, Finance and Management, 2*(1), 19–39.

Coakley, J. R., & Brown, C. E. (2000). Artificial neural networks in accounting and finance: Modeling issues. *Intelligent Systems in Accounting, Finance and Management, 9*(2), 119–144.

Coakley, J. R., & Loebbecke, J. K. (1985). The expectation of accounting errors in medium-sized manufacturing firms. *Advances in Accounting, 2*(1985), 199–246.

Coakley, J., McFarlane, D., & Perley, W. (1992, April). *Alternative criteria for evaluating artificial neural network performance.* Paper presented at the TIMS/ORSA Joint National Meeting, Orlando, FL.

Cogger, K. O. (1981). A time-series analytic approach to aggregation issues in accounting data. *Journal of Accounting Research*, *19*(2), 285–298.

Colbert, J. L. (1994). Analytical procedures for management accountants and auditors. *Managerial Auditing Journal, 9*(5), 3–7.

Colbert, J. L., & Wilson, A. (1991). The audit strategy: Analytical procedures as substantive tests. *National Public Accountant, 36*, 38–41.

Curry, B., & Peel, M. J. (1998). Neural networks and business forecasting: An application to cross-sectional audit fee data. *International Journal of Commerce and Management, 8*(2), 94–120.

Cybenko, G. (1988). *Continuous valued neural networks with two hidden layers are sufficient*. Tufts University: Medford, MA.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems, 2*(4), 303–314.

Davis, J. T., Massey, A. P., & Lovell, R. E., II. (1997). Supporting a complex audit judgment task: An expert network approach. *European Journal of Operational Research*, *103*(2), 350–372.

DBS. (2017). China dairy: Downstream is key. Retrieved from https://www.dbs.com/aics/templatedata/article/generic/data/en/GR/112017/171130_insights_china_dairy_downstream_is_key.xml

Diamantaras, K. I., & Kung, S. Y. (1996). *Principal component neural networks: Theory and applications*. New York, NY: John Wiley & Sons.

Dong, W., Liao, S. S., Fang, B., Cheng, X., Chen, Z., & Fan, W. (2014, June). *The detection of fraudulent financial statements: An integrated language model.* Paper presented at the Pacific Asia Conference on Information Systems, Chengdu, China.

Dugan, M. T., Gentry, J. A., & Shriver, K. A. (1985). The x-11 model – A new analytical review technique for the auditor. *Auditing: A Journal of Practice & Theory, 4*(2), 11–22.

Dzeng, S. C. (1994). A comparison of analytical procedure expectation models using both aggregate and disaggregate data. *Auditing: A Journal of Practice & Theory, 13*(2), 1–24.

Elliott, R. K. (1983). Unique audit methods: Peat Marwick International. *Auditing: A Journal of Practice & Theory, 2*(2), 1–12.

Etheridge, H. L., Sriram, R. S., & Hsu, H. (2000). A comparison of selected artificial neural networks that help auditors evaluate client financial viability. *Decision Sciences, 31*(2), 531–550.

Fahlman, S. E., & Lebiere, C. (1990). The cascade-correlation learning architecture. In D. S. Touretzky (Ed.), *Advances in neural information processing systems* (Vol. 2, pp. 524–532). Retrieved from https://papers.nips.cc/paper/207-the-cascade-correlation-learning-architecture.pdf

Fanning, K. M., & Cogger, K. O. (1994). A comparative analysis of artificial neural networks using financial distress prediction. *Intelligent Systems in Accounting, Finance and Management, 3*(4), 241–252.

Fanning, K. M., & Cogger, K. O. (1998). Neural network detection of management fraud using published financial data. *International Journal of Intelligent Systems in Accounting, Finance and Management, 7*(1), 21–41.

Feroz, E. H., Kwon, T. M., Pastena, V. S., & Park, K. (2000). The efficacy of red flags in predicting the SEC's targets: An artificial neural networks approach. *International Journal of Intelligent Systems in Accounting, Finance and Management, 9*(3), 145–157.

Fisher, B. (1999). Mellon creates FraudWatch to predict and manage risk using neural technology. *Journal of Retail Banking Services, 21*(1), 15–18.

Fraser, I. A., Hatherly, D. J., & Lin, K. Z. (1997). An empirical investigation of the use of analytical review by external auditors. *The British Accounting Review, 29*(1), 35–47.

Glancy, F. H., & Yadav, S. B. (2011). A computational model for financial reporting fraud detection. *Decision Support Systems, 50*(3), 595–601.

Glover, S. M., Prawitt, D. F., & Drake, M. S. (2014). Between a rock and a hard place: A path forward for using substantive analytical procedures in auditing large P&L accounts: Commentary and analysis. *Auditing: A Journal of Practice & Theory, 34*(3), 161–179.

Glover, S. M., Prawitt, D. F., & Wilks, T. J. (2005). Why do auditors over-rely on weak analytical procedures? The role of outcome and precision. *Auditing: A Journal of Practice & Theory, 24*(Suppl. 1), 197–220.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1). Cambridge, MA: MIT Press.

Green, B. P., & Calderon, T. G. (1994). Using real-world cases to illustrate the power of analytical procedures. *Journal of Accounting Education, 12*(3), 245–268.

Green, B. P., & Choi, J. H. (1997). Assessing the risk of management fraud through neural network technology. *Auditing: A Journal of Practice & Theory, 16*(1), 14–28.

Hansen, J. V., McDonald, J. B., & Stice, J. D. (1992). Artificial intelligence and generalized qualitative-response models: An empirical test on two audit decision-making domains. *Decision Sciences, 23*(3), 708–723.

Hecht-Nielsen, R. (1989). Neurocomputer applications. In R. Eckmiller & C. v.d. Malsburg (Eds.), *Neural computers* (pp. 445–453). Berlin, Germany: Springer.

Hirst, D., & Koonce, L. (1996). Audit analytical procedures: A field investigation. *Contemporary Accounting Research, 13*(2), 457–486.

Hoitash, R., Kogan, A., & Vasarhelyi, M. A. (2006). Peer-based approach for analytical procedures. *Auditing: A Journal of Practice & Theory, 25*(2), 53–84.

Holstrum, G. L., & Messier, W. F. (1982). A review and integration of empirical research on materiality. *Auditing: A Journal of Practice & Theory, 2*(1), 45–63.

Hoogs, B., Kiehl, T., Lacomb, C., & Senturk, D. (2007). A genetic algorithm approach to detecting temporal patterns indicative of financial statement fraud. *Intelligent Systems in Accounting, Finance and Management, 15*(1–2), 41–56.

Huang, S. Y. (2013). Fraud detection model by using support vector machine techniques. *International Journal of Digital Content Technology and its Applications, 7*(2), 32–42.

Humpherys, S. L., Moffitt, K. C., Burns, M. B., Burgoon, J. K., & Felix, W. F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems, 50*(3), 585–594.

Hylas, R. E., & Ashton, R. H. (1982). Audit detection of financial statement errors. *Accounting Review*, *57,* 751–765.

Institute of Chartered Accountants in England and Wales (ICAEW). (2016). Data analytics for external auditors. Retrieved from https://www.icaew.com/-/media/corporate/files/technical/iaa/tecpln14726-iaae-data-analytics---web-version.ashx

Icerman, R. C., & Hillison, W. A. (1991). Disposition of audit-detected errors – Some evidence on evaluative materiality. *Auditing: A Journal of Practice & Theory*, *10*(1), 22–34.

International Standard on Auditing (ISA) 315. (2009). Identifying and assessing the risks of material misstatement through understanding the entity and its environment. Retrieved from http://www.ifac.org/system/files/downloads/a017-2010-iaasb-handbook-isa-315.pdf

International Standard on Auditing (ISA) 520. (2009). Analytical procedure. Retrieved from http://www.ifac.org/system/files/downloads/a026-2010-iaasb-handbook-isa-520.pdf

Issa, H., Sun, T., & Vasarhelyi, M. A. (2016). Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation. *Journal of Emerging Technologies in Accounting, 13*(2), 1–20.

Jhee, W. C., & Lee, J. K. (1993). Performance of neural networks in managerial forecasting. *Intelligent Systems in Accounting, Finance and Management, 2*(1), 55–71.

Kinney, W., Jr. (1978). ARIMA and regression in analytical review: An empirical test. *Accounting Review*, 53(1), 48–60.

Kinney, W., Jr. (1979). The predictive power of limited information in preliminary analytical review: An empirical study. *Journal of Accounting Research*, *17*(Suppl.), 148–165.

Kinney, W., Jr. (1987). Attention directing analytical review using accounting ratios: A case study. *Auditing: A Journal of Practice & Theory, 6*(2), 59–73.

Kinney, W., Jr., McDaniel, L., & Martin, R. (2005). *Sequential quantification guidance to assist auditors' subjective assessments in substantive analytical procedures* (Working paper). University of Kentucky: Lexington, KY.

Kinney, W., Jr., & Salamon, G. L. (1982). Regression analysis in auditing: A comparison of alternative investigation rules. *Journal of Accounting Research*, *20*(2), 350–366.

Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications, 32*(4), 995–1003.

Knechel, W. R. (1986). Applications and implementation a simulation study of the relative effectiveness of alternative analytical review procedures. *Decision Sciences, 17*(3), 376–394.

Knechel, W. R. (1988). The effectiveness of statistical analytical review as a substantive auditing procedure: A simulation analysis. *Accounting Review*, *63*(1), 74–95.

Kogan, A., Alles, M. G., Vasarhelyi, M. A., & Wu, J. (2010). *Analytical procedures for continuous data level auditing: Continuity equations*. Retrieved from http://raw.rutgers.edu/docs/Innovations/Continuity%20Equations.pdf

Koh, H. C., & Tan, S. S. (1999). A neural network approach to the prediction of going concern status. *Accounting and Business Research, 29*(3), 211–216.

Koskivaara, E. (2003). *Artificial neural networks in auditing: State of the art* (TUCS Technical Reports 509). Turku, Finland: Turku Centre for Computer Science.

Kou, Y., Lu, C.-T., Sirwongwattana, S., & Huang, Y.-P. (2004, March). *Survey of fraud detection techniques.* Paper presented at the International Conference on Networking, Sensing and Control, Taipei, Taiwan.

Lapedes, A., & Farber, R. (1987). *Nonlinear signal processing using neural networks: Prediction and system modelling*. Retrieved from http://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-87-2662

Lenard, M. J., Alam, P., & Madey, G. R. (1995). The application of neural networks and a qualitative response model to the auditor's going concern uncertainty decision. *Decision Sciences, 26*(2), 209–227.

Lev, B. (1980). On the use of index models in analytical reviews by auditors. *Journal of Accounting Research*, *18*(2), 524–550.

Libby, R., & Luft, J. (1993). Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. *Accounting, Organizations and Society, 18*(5), 425–450.

Lin, K. Z., & Fraser, I. A. (2003). The use of analytical procedures by external auditors in Canada. *Journal of International Accounting, Auditing and Taxation, 12*(2), 153–168.

Lin, K. Z., Fraser, I. A., & Hatherly, D. J. (2000). An experimental study of auditor analytical review judgements. *Journal of Business Finance & Accounting, 27*(7–8), 821–857.

Loebbecke, J. K., & Steinbart, P. J. (1987). An investigation of the use of preliminary analytical review to provide substantive audit evidence. *Auditing: A Journal of Practice & Theory, 6*(2), 74–89.

Loebbecke, J. K., & Willingham, J. (1988). Review of SEC accounting and auditing enforcement releases. Unpublished manuscript.

Lorek, K. S., Wheeler, S., Icerman, R., & Fordham, D. (1995). An investigation of the feasibility of using statistically-based models as analytical procedures. *Advances in Accounting, 13*, 87–110.

Masters, T., & Schwartz, M. (1994). Practical neural network recipes in C. *IEEE Transactions on Neural Networks, 5*(5), 853–853.

McDaniel, L. S., & Simmons, L. E. (2007). Auditors' assessment and incorporation of expectation precision in evidential analytical procedures. *Auditing: A Journal of Practice & Theory, 26*(1), 1–18.

McKee, T. E. (1989). *Modern analytical auditing: Practical guidance for auditors and accountants*. New York, NY: Quorum Books.

Moreira, M., & Fiesler, E. (1995). *Neural networks with adaptive learning rate and momentum terms*. Retrieved from https://infoscience.epfl.ch/record/82307/files/95-04.pdf

Mulqueen, J. (1996). Neural nets block fraud. *CommunicationWeek, 625*, 95–96.

Neter, J. (1980). *Two case studies on use of regression for analytic review.* Paper presented at the Symposium on Auditing Research IV, Urbana-Champaign, IL.

Ngai, E., Hu, Y., Wong, Y., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems, 50*(3), 559–569.

Odom, M. D., & Sharda, R. (1990, June). *A neural network model for bankruptcy prediction.* Paper presented at the International Joint Conference on Neural Networks, San Diego, CA.

Pany, K., & Wheeler, S. (1992). A test of analytical procedure effectiveness. *The CPA Journal, 62*(6), 64.

Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. Retrieved from https://arxiv.org/ftp/arxiv/papers/1009/1009.6119.pdf

Ramamoorti, S., Bailey, J., & Traver, R. O. (1999). Risk assessment in internal auditing: A neural network approach. *International Journal of Intelligent Systems in Accounting, Finance and Management, 8*(3), 159–180.

Ravisankar, P., Ravi, V., Rao, G. R., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems, 50*(2), 491–500.

Reed, R. (1993). Pruning algorithms – A survey. *IEEE Transactions on Neural Networks, 4*(5), 740–747.

Salchenberger, L. M., Cinar, E., & Lash, N. A. (1992). Neural networks: A new tool for predicting thrift failures. *Decision Sciences, 23*(4), 899–916.

Sarma, K. S. (2013). *Predictive modeling with SAS Enterprise Miner: Practical solutions for business applications*. Cary, NC: SAS Institute.

Shanker, M., Hu, M. Y., & Hung, M. S. (1996). Effect of data standardization on neural network training. *Omega, 24*(4), 385–397.

Smith, G., Psaros, J., & Holmes, S. (1999). A research note on the use and perceived usefulness of analytical procedures by Australian auditors. *Australian Accounting Review, 9*(18), 64–72.

Stringer, K. W., & Stewart, T. R. (1986). *Statistical techniques for analytical review in auditing*. New York, NY: John Wiley & Sons.

Subramanian, V., Hung, M. S., & Hu, M. Y. (1993). An experimental evaluation of neural networks for classification. *Computers & Operations Research, 20*(7), 769–782.

Surkan, A. J., & Singleton, J. C. (1990). *Neural networks for bond rating improved by multiple hidden layers.* Paper presented at the International Joint Conference on Neural Networks, San Diego, CA.

Tabor, R. H., & Willis, J. T. (1985). Empirical evidence on the changing role of analytical review procedures. *Auditing: A Journal of Practice & Theory, 4*(2), 93–109.

Trompeter, G., & Wright, A. (2010). The world has changed—Have analytical procedure practices? *Contemporary Accounting Research, 27*(2), 669–700.

Turban, E., Aronson, J. E., Liang, T. P., & Sharda, R. (2007). *Decision support and business intelligence system* (8th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Vellido, A., Lisboa, P. J., & Vaughan, J. (1999). Neural networks in business: A survey of applications (1992–1998). *Expert Systems with Applications, 17*(1), 51–70.

Wessels, L. F., & Barnard, E. (1992). Avoiding false local minima by proper initialization of connections. *IEEE Transactions on Neural Networks, 3*(6), 899–905.

West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security, 57*, 47–66.

Wheeler, S., & Pany, K. (1990). Assessing the performance of analytical procedures: A best case scenario. *Accounting Review*, *65*(3), 557–577.

Wild, J. J. (1987). The prediction performance of a structural model of accounting numbers. *Journal of Accounting Research*, 25(1), 139–160.

Wilson, A. C. (1991). Use of regression models as analytical procedures: An empirical investigation of effect of data dispersion on auditor decisions. *Journal of Accounting, Auditing & Finance, 6*(3), 365–381.

Wilson, A. C., & Colbert, J. (1989). An analysis of simple and rigorous decision models as analytical procedures. *Accounting Horizons, 3*(4), 79–83.

Wilson, A. C., Colbert, J. L., & Minyard, D. H. (1991). ARIMA as an analytical procedure: Relative efficiency and effectiveness of a limited number of observations. *Review of Financial Economics, 26*(2), 1.

Wong, B. K., Bodnovich, T. A., & Selvi, Y. (1997). Neural network applications in business: A review and analysis of the literature (1988–1995). *Decision Support Systems, 19*(4), 301–320.

Wong, B. K., & Selvi, Y. (1998). Neural network applications in finance: A review and analysis of literature (1990–1996). *Information & Management, 34*(3), 129–139.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting, 14*(1), 35–62.

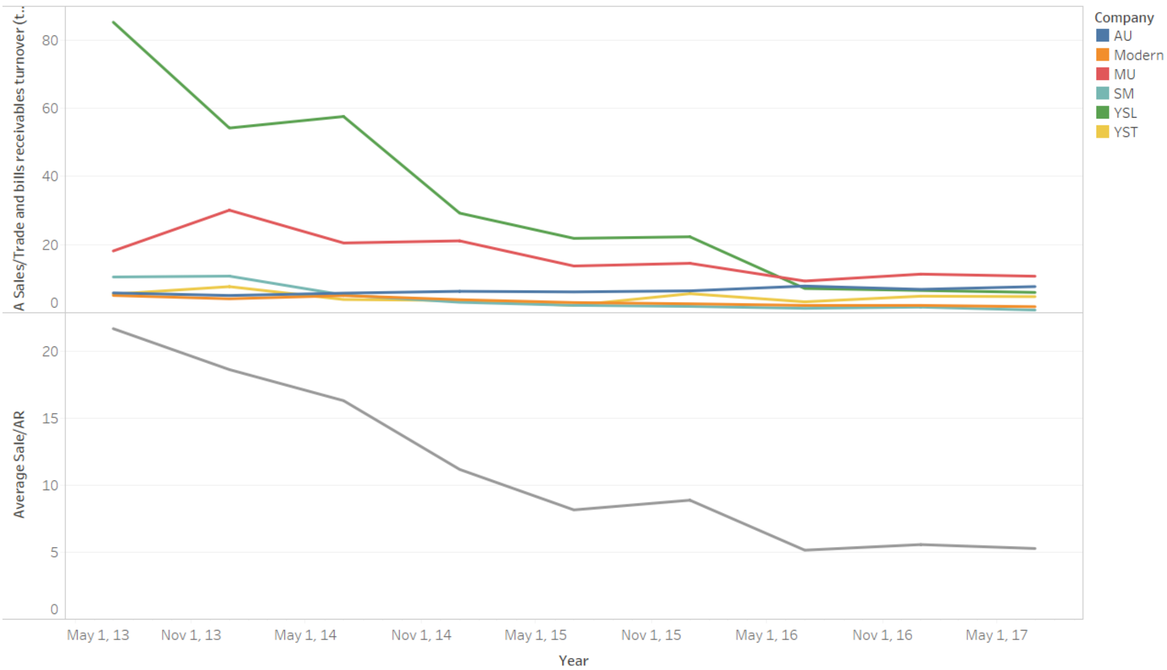
APPENDIX I – LIST OF CASE FIRMS

s/n	Firm	Founded	HQ	Ticker	Products
1	Ausnutria Dairy (AU)	2003	Hong Kong	1717	Infant formula
2	Yashili International (YSL)	1983	Guangdong	1230	Infant formula
3	China Mengniu Dairy (MU)	1999	Inner Mongolia	2319	Liquid milk etc.
4	China Modern Dairy (Modern)	2005	Anhui	1117	Dairy milk etc.
5	China Shengmu Organic (SM)	2009	Inner Mongolia	1432	Dairy milk etc.
6	Yuanshengtai Dairy (YST)	2008	Heilongjiang	1431	Dairy milk
7	China Huishan Dairy (HS)	1951	Liaoning	6863	Dairy milk etc.
8	China Zhongdi Dairy	2002	Beijing	1492	Dairy milk
9	Daqing Dairy	1970	Heilongjiang	1007	Dairy milk etc.
10	Lanzhou Zhuangyuan	2000	Gansu	1533	Dairy milk etc.
11	CIMC Enric Holdings	2004	Guangdong	3899	Machinery (Dairy etc.)
12	Four Seas Mercantile	1971	Hong Kong	0374	Processed food
13	Sinopharm International	2003	Shanghai	1099	Beverages
14	Uni-President China	1992	Shanghai	0220	Beverages
15	Ping An Insurance	1988	Hong Kong	2318	Finance (Dairy)

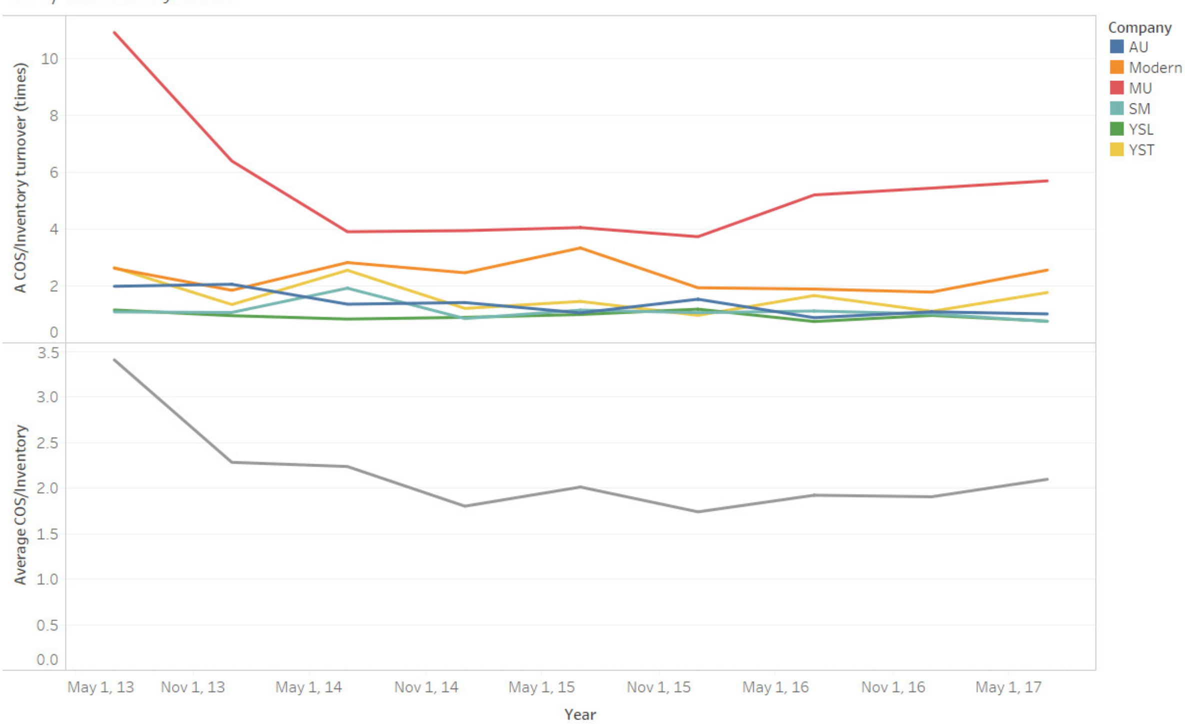
Source: http://www.chinaag.org/hkg-listed/

APPENDIX II – COMPARISON BETWEEN FIRMS' AND AVERAGED RATIOS

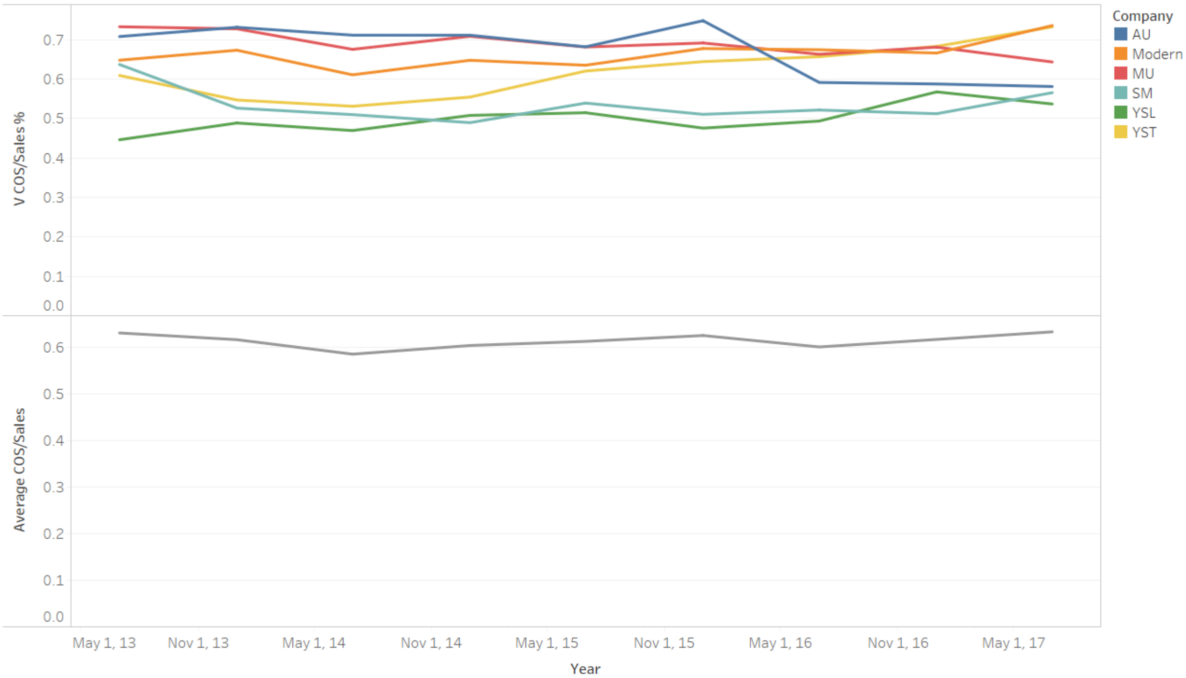
Sales/AR ratio



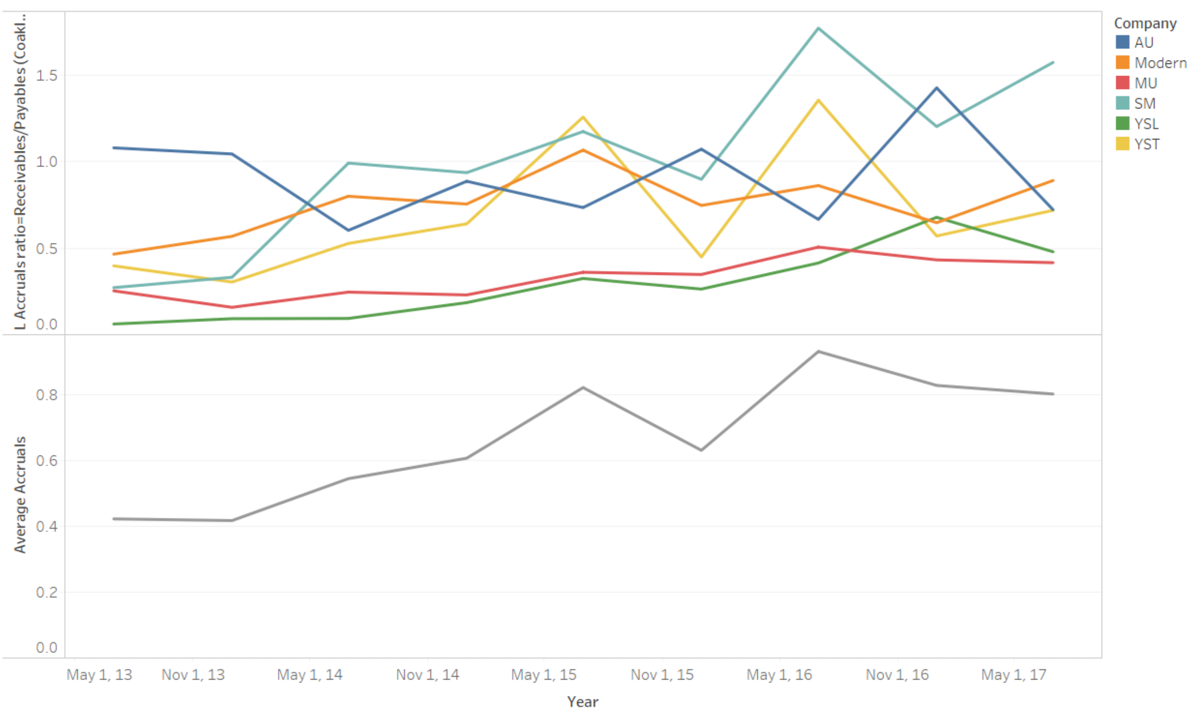
COS/Inventory ratio



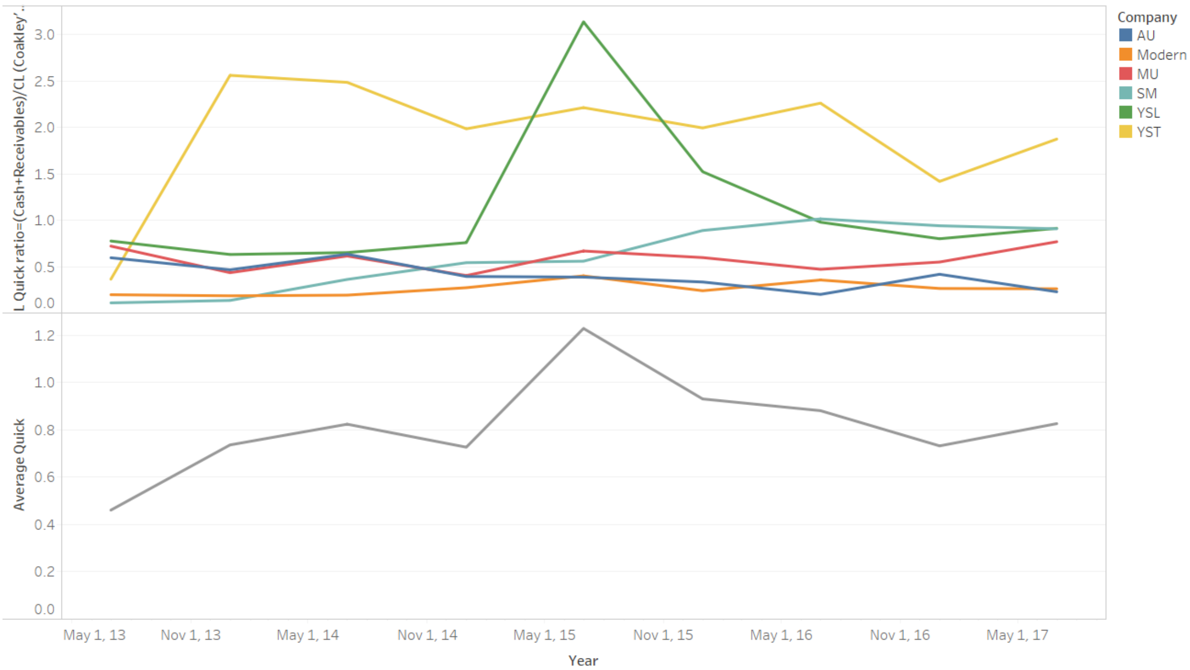
COS/Sales ratio



Accruals ratio



Quick ratio



Company
AU
Modern
MU
SM
YSL
YST
L Quick ratio=(Cash+Receivables)/CL (Coakley'..
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Average Quick
1.2
1.0
0.8
0.6
0.4
0.2
0.0
May 1, 13
Nov 1, 13
May 1, 14
Nov 1, 14
May 1, 15
Nov 1, 15
May 1, 16
Nov 1, 16
May 1, 17
Year

APPENDIX III – THE USE OF 15, 20 AND 30 TIMES M

Effect of error size on Type II error rate in terms of alpha risk levels and sources of error

Alpha risk/Source	Financial ratio	Regression	Neural network
0.10 – NA	0.86	0.64	0.43
0.33 – NA	0.76	0.54	0.33
0.10 – ADJ	0.79	0.64	0.45
0.10 – ADJ	0.68	0.52	0.36
Purchases – NA	0.80	0.58	0.38
Sales – NA	0.82	0.60	0.39
Purchases – ADJ	0.70	0.56	0.40
Sales – ADJ	0.78	0.60	0.40

APPENDIX IV – PARAMETERS OF THE NEURAL NETWORK MODEL

Network in SAS Enterprise Miner 14.3	
Architecture	Multilayer Perceptron
Direct Connection	No
Input Variables	22 (Input data stream=43)
Number of Hidden Layer	1
Number of Hidden Units	11
Output Variables	7
Randomization Distribution	Normal
Randomization Center	0.0
Randomization Scale	0.1
Input Standardization	Standard Deviation
Hidden Layer Combination Function	Linear
Hidden Layer Activation Function	Hyperbolic Tangent
Hidden Bias	Yes
Target Layer Combination Function	Linear
Target Layer Activation Function	Identity
Target Layer Error Function	Normal (SSE)
Target Bias	Yes
Weight Decay	0.1
Termination	Convergence, Validation error, Maximum iterations
Model selection criterion property	Average error
Optimization in SAS Enterprise Miner 14.3	
Learning method	Supervised
Training Technique	Feedforward gradient descent back-propagation
Maximum Iterations	100
Maximum Time	4 Hours
Nonlinear Option	Default
Learn	0.001
Maximum Learning	50.0
Minimum Learning	1.0E-5
Momentum	0.001
Maximum Momentum	1.75
Tilt	0.0
Autotune Hyperparameters in SAS Visual Data Mining and Machine Learning	
Optimization algorithm	Stochastic gradient descent SGP
Number of Hidden Layers	0
Number of Hidden Neurons	0
Regularization L1	0.02726829
Regularization L2	0.05562433
Learning Rate	0.01709667
Annealing Rate	0.0000000001